Credenza, revisione e razionalità

Una ricostruzione formale della teoria AGM e dei suoi fondamenti epistemici

Movimento I — Il problema e la sua genealogia

Che cosa vuol dire cambiare idea in modo razionale? La domanda, nella sua apparente semplicità, nasconde una delle questioni più intricate che la filosofia analitica abbia ereditato dall'epistemologia classica. Non si tratta soltanto di descrivere come gli agenti cognitivi di fatto modifichino le proprie credenze quando ricevono informazioni nuove — compito che apparterrebbe alla psicologia o alla scienza cognitiva — bensì di stabilire quali vincoli normativi debbano soddisfare tali modificazioni perché possano dirsi *razionali*. La distinzione tra descrizione e norma è qui essenziale: una teoria della revisione doxastica è, prima di tutto, una teoria della coerenza pratica della mente che conosce.

Il punto di partenza naturale è la concezione classica della credenza come appartenenza a un insieme. Se un agente detiene un corpus doxastico K — inteso come l'insieme di tutte le proposizioni che egli crede vere — allora la sua posizione epistemica è interamente catturata dalla struttura algebrica di K. Questa astrazione è volutamente idealizzante: non si presuppone che K debba essere finito, né che l'agente sia psicologicamente capace di calcolare tutte le conseguenze di ciò che crede; si presuppone soltanto che K sia *logicamente chiuso*, nel senso che se $\alpha \in K$ e $\alpha \vdash \beta$, allora $\beta \in K$. Formalmente, K è ciò che la logica chiama una *teoria*: un insieme di enunciati chiuso sotto conseguenza logica per una data logica di sfondo L.

Questa idealizzazione, che Gärdenfors chiamerà l'*assunzione della chiusura logica*, è già filosoficamente carica. Essa esclude dal novero degli agenti razionali qualunque mente finita che non sia in grado di dedurre tutte le implicazioni delle proprie credenze, il che, dal punto di vista cognitivo, esclude ogni mente reale. Tuttavia la teoria AGM non ambisce a essere una psicologia; è piuttosto una *logica* della revisione, nel senso in cui la logica deontica non descrive come gli esseri umani ragionino circa obblighi e permessi, bensì caratterizza le relazioni strutturali tra tali concetti. La critica alla mancata fedeltà cognitiva è quindi un colpo a vuoto, almeno nel merito della struttura originaria della proposta.

Il nucleo del problema si presenta non appena si chiedono le condizioni in cui K deve mutare. Tre operazioni fondamentali vengono distinte con precisione chirurgica: l'*espansione*, la *contrazione* e la *revisione*. Nella prima, un agente aggiunge una nuova

credenza α al proprio corpus senza rimuovere nulla di ciò che già credeva — si scrive $K + \alpha$ e la sua definizione è banale: $K + \alpha = Cn(K \cup \{\alpha\})$, dove Cn denota la chiusura per conseguenza. Nella seconda, l'agente rinuncia a una credenza α che già possedeva, producendo un corpus $K - \alpha$ che non contenga più α tra le sue conseguenze. Nella terza — la più filosoficamente impegnativa — l'agente è costretto ad accogliere una credenza α che contraddice qualcosa che già credeva, producendo un corpus $K * \alpha$ che contenga α ma resti consistente.

La contrazione e la revisione sono, come Alchourrón, Gärdenfors e Makinson mostrarono nel loro saggio del 1985 in *The Journal of Symbolic Logic*, profondamente connesse attraverso l'identità di Levi e la sua inversa. L'identità di Levi afferma che la revisione per α può essere costruita come espansione con α dopo aver contratto per $\neg\alpha$:

$$K * \alpha = (K - \neg\alpha) + \alpha$$

L'identità di Harper, nella direzione opposta, esprime la contrazione come intersezione tra il corpus originario e la revisione per la negazione:

$$K - \alpha = K \cap (K * \neg\alpha)$$

Queste due identità rivelano che, a livello strutturale, le operazioni non sono indipendenti: basta fornire una semantica soddisfacente per l'una per ricavare l'altra. Nel seguito si preferirà lavorare sulla revisione, per ragioni di eleganza e perché è su di essa che la letteratura successiva ha prodotto i contributi più ricchi, ma ogni risultato ha il proprio duale per la contrazione.

Movimento II — I postulati e la loro giustificazione

Il cuore della proposta AGM è un sistema di otto postulati per la revisione, tradizionalmente divisi in sei *postulati di base* e due *postulati supplementari*. Essi non vengono presentati come assiomi senza motivazione, bensì come esplicitazioni di condizioni di razionalità che un agente riflessivo dovrebbe riconoscere come vincolanti. Questo stile di giustificazione — per introspection dell'agente ideale piuttosto che per dimostrazione formale — è filosoficamente caratteristico della tradizione che va da Kant ai pragmatisti e che in logica trova espressione nel metodo delle *intuizioni pompe*. Vediamo i postulati uno per uno, con la loro motivazione.

Il primo postulato, detto di *chiusura*, richiede che il corpo revisionato $K * \alpha$ sia anch'esso una teoria:

$$(K*1)\quad K * \alpha = Cn(K * \alpha)$$

La motivazione è immediata: se la revisione produce un insieme non chiuso, l'agente si trova in una posizione epistemica incoerente in cui detiene alcune credenze ma non le

loro conseguenze ovvie. Questo sarebbe una forma di irrazionalità procedurale difficile da difendere.

Il secondo postulato, di *successo*, garantisce che la nuova informazione venga effettivamente incorporata:

(K*2) $\alpha \in K * \alpha$

Senza questo requisito, la revisione potrebbe in linea di principio ignorare l'evidenza che l'ha provocata, il che renderebbe il concetto stesso di revisione vuoto.

Il terzo postulato, di *inclusione*, impone che la revisione non produca credenze del tutto nuove rispetto all'unione di K con {α}:

(K*3) $K * \alpha \subseteq K + \alpha = Cn(K \cup \{\alpha\})$

Questo cattura l'idea che la revisione sia *conservativa*: l'agente non inventa credenze dal nulla, ma si limita a risistemare quelle che già possedeva alla luce della nuova informazione. La revisione non è creatività epistemica; è riarrangiamento.

Il quarto postulato, di *vacuità*, specifica cosa accade quando la nuova informazione è compatibile con il corpus preesistente:

(K*4) Se $\neg\alpha \notin K$, allora $K + \alpha \subseteq K * \alpha$

Se α non contraddice nulla di ciò che l'agente già crede, la revisione si riduce all'espansione: non c'è nulla da rimuovere, quindi il corpus revisionato deve contenere almeno tutto ciò che l'espansione produrrebbe. Insieme a (K*3), questo implica che quando $\neg\alpha \notin K$, vale $K * \alpha = K + \alpha$ — la revisione collassa sull'espansione nei casi non conflittuali.

Il quinto postulato, di *consistenza*, richiede che il risultato della revisione sia consistente, eccetto nel caso degenere in cui α sia essa stessa una contraddizione:

(K*5) $K * \alpha = K_\bot$ solo se $\vdash \neg\alpha$

dove $K_\bot$ denota la teoria inconsistente, ovvero l'insieme di tutti gli enunciati. In tutte le situazioni non degeneri, la revisione deve preservare la consistenza del corpus. Questo è il postulato che rende la revisione genuinamente diversa dalla semplice espansione: quando α contraddice qualcosa in K, bisogna rimuovere qualcosa da K prima di aggiungere α, e la rimozione deve essere sufficiente a ristabilire la coerenza.

Il sesto postulato, di *estensionalità*, afferma che logicamente equivalenti input producono identici output:

(K*6) Se $\vdash \alpha \leftrightarrow \beta$, allora $K * \alpha = K * \beta$

Questo è un requisito di indifferenza alla forma sintattica: la revisione deve essere sensibile al contenuto logico dell'informazione, non alla sua veste linguistica. Un

agente che revisioni diversamente a seconda di come venga formulata la stessa proposizione sarebbe vittima di una fallacia puramente sintattica.

I due postulati supplementari, (K*7) e (K*8), governano la coerenza della revisione quando si forniscono informazioni *congiunte*. Il settimo afferma:

(K*7) $K * (\alpha \wedge \beta) \subseteq (K * \alpha) + \beta$

Il che significa che revisionare per $\alpha \wedge \beta$ deve produrre qualcosa contenuto nell'espansione di $K * \alpha$ con β. Intuitivamente: se prima si revisiona per α e poi si espande con β, il risultato deve essere almeno tanto ricco quanto il revisore congiunto. L'ottavo postulato aggiunge la condizione simmetrica:

(K*8) Se $\neg\beta \notin K * \alpha$, allora $(K * \alpha) + \beta \subseteq K * (\alpha \wedge \beta)$

Ovvero: se β non contraddice il corpus revisionato $K * \alpha$, allora la revisione congiunta deve contenere almeno quanto l'espansione. Insieme, (K*7) e (K*8) implicano che quando $\neg\beta \notin K * \alpha$:

$K * (\alpha \wedge \beta) = (K * \alpha) + \beta$

Questo risultato ha un nome: è il *principio di minima mutilazione congiunta*, e afferma che la revisione per una congiunzione, quando la seconda congiunzione non disturba il corpus già revisionato per la prima, equivale a revisionare per la prima congiunzione e poi espandere con la seconda. La struttura è quella di una revisione iterata senza conflitti aggiuntivi.

La questione filosofica più sottile riguarda la giustificazione di (K*7) e (K*8). Mentre i primi sei postulati godono di motivazioni abbastanza trasparenti, i supplementari sono stati oggetti di controversia. Peter Gärdenfors stesso, in lavori successivi al 1985, ha riconosciuto che la loro giustificazione intuitiva è meno immediata, e che essi assumono implicitamente una proprietà strutturale degli ordinamenti dossastici che merita trattamento esplicito. Tale trattamento giunse attraverso la semantica delle sfere di Adam Grove.

Movimento III — La semantica delle sfere di Grove

La dimostrazione che i postulati AGM sono non soltanto intuitivamente plausibili ma anche formalmente caratterizzati da una classe precisa di strutture semantiche fu offerta da Adam Grove nel 1988, in un articolo rimasto tra i più eleganti della logica epistemica del decennio. L'idea portante di Grove è che l'insieme dei *mondi possibili* — qui intesi come assegnazioni di valori di verità a tutti gli enunciati del linguaggio — possa essere organizzato in una struttura concentrica di *sfere* attorno al corpus doxastico dell'agente, e che questa struttura catturi esattamente la nozione intuitiva di *vicinanza epistemica*: alcuni mondi, pur non essendo compatibili con ciò che l'agente crede, sono comunque più simili alla sua visione del mondo di altri.

Formalmente, sia W l'insieme di tutti i modelli (o mondi possibili) per il linguaggio L. Un *sistema di sfere* centrato su un insieme $K \subseteq W$ è una famiglia S di sottoinsiemi di W tale che:

(S1) S è totalmente ordinata da $\subseteq$ (cioè per ogni $A, B \in S$: $A \subseteq B$ oppure $B \subseteq A$)

(S2) $[K] \in S$, dove $[K] = \{w \in W \mid w \vDash \alpha \text{ per ogni } \alpha \in K\}$ è l'insieme dei mondi compatibili con K

(S3) $[K]$ è la sfera minima (il centro del sistema)

(S4) $W \in S$ (la sfera massima è l'intero spazio dei mondi)

La struttura è quella di una famiglia di insiemi concentrici, dal nucleo $[K]$ — i mondi più simili a ciò che l'agente crede — fino all'intero W. L'ordinamento totale garantisce che non vi siano sfere incomparabili: per ogni due mondi w_1 e w_2, uno dei due appartiene a tutte le sfere a cui appartiene l'altro, oppure viceversa. Questo equivale a dire che la relazione di preordine totale $\preccurlyeq$ su W definita da:

$w_1 \preccurlyeq w_2$ sse per ogni $A \in S$: se $w_1 \in A$ allora $w_2 \in A$

è un preordine totale (riflessivo, transitivo e connesso). In questo preordine, $w_1 \preccurlyeq w_2$ si legge "w_1 è almeno altrettanto plausibile quanto w_2" o, equivalentemente, "w_1 è almeno altrettanto prossimo al corpus K quanto w_2".

La revisione viene ora definita semanticamente. Dato un enunciato α, sia $[\alpha] = \{w \in W \mid w \vDash \alpha\}$ l'insieme dei mondi che rendono vera α. La revisione di K per α seleziona i mondi di $[\alpha]$ che appartengono alla sfera più piccola (più interna) che interseca $[\alpha]$:

$K * \alpha$ è l'insieme di tutti gli enunciati veri in ogni $w \in (S_\alpha \cap [\alpha])$

dove $S_\alpha = \bigcap\{A \in S \mid A \cap [\alpha] \neq \varnothing\}$ è la sfera minima di S che ha qualche elemento in comune con $[\alpha]$. In simboli, identificando le teorie con gli insiemi di mondi che le soddisfano:

$$[K * \alpha] = S_\alpha \cap [\alpha]$$

Questa definizione ha un'interpretazione luminosa: per revisionare alla luce di α, si prendono i mondi che rendono vera α e che sono, secondo l'ordinamento di plausibilità, il più vicini possibile al corpus originario. È una forma di ottimizzazione: si minimizza la distanza dalla propria visione del mondo soggetta al vincolo che la nuova informazione sia vera.

Il *teorema di rappresentazione di Grove* afferma che:

Teorema (Grove, 1988). Un'operazione $* : 2^L \times L \to 2^L$ soddisfa i postulati (K*1)–(K*8) se e soltanto se esiste un sistema di sfere S centrato su $[K]$ tale che $[K * \alpha] = S_\alpha \cap [\alpha]$ per ogni α consistente.

La dimostrazione procede in due direzioni. La direzione **da sfere a postulati** — mostrare che ogni operazione di revisione definita da un sistema di sfere soddisfa tutti e otto i postulati — è relativamente meccanica, anche se richiede attenzione. Verifichiamo i casi meno ovvi.

Per (K*5): se α è consistente, allora $[\alpha] \neq \varnothing$, dunque $S_\alpha \cap [\alpha] \neq \varnothing$ (perché S_α, come unione di sfere, contiene W, e quindi interseca ogni insieme non vuoto nel suo raggio), e quindi $K * \alpha$ è consistente. ✓

Per (K*7) e (K*8): sia S_α la sfera minima che interseca $[\alpha]$, e sia $S_{\alpha \wedge \beta}$ la sfera minima che interseca $[\alpha \wedge \beta] = [\alpha] \cap [\beta]$. Poiché $[\alpha \wedge \beta] \subseteq [\alpha]$, ogni sfera che interseca $[\alpha \wedge \beta]$ interseca anche $[\alpha]$, dunque $S_\alpha \subseteq S_{\alpha \wedge \beta}$. Ora:

$[K * (\alpha \wedge \beta)] = S_{\alpha \wedge \beta} \cap [\alpha] \cap [\beta]$

$[(K * \alpha) + \beta] = Cn([K * \alpha] \cup \{\beta\})$, che semanticamente corrisponde a $[K * \alpha] \cap [\beta] = (S_\alpha \cap [\alpha]) \cap [\beta]$

Per (K*7) si deve mostrare che $S_{\alpha \wedge \beta} \cap [\alpha] \cap [\beta] \subseteq S_\alpha \cap [\alpha] \cap [\beta]$, ovvero $S_{\alpha \wedge \beta} \subseteq S_\alpha$. Poiché ogni sfera che interseca $[\alpha \wedge \beta]$ interseca $[\alpha]$, la sfera minima per α è contenuta nella sfera minima per $\alpha \wedge \beta$: $S_\alpha \subseteq S_{\alpha \wedge \beta}$. Questo dimostra la *direzione opposta*, dunque (K*7) richiede $S_{\alpha \wedge \beta} \cap [\alpha \cap \beta] \subseteq S_\alpha \cap [\alpha] \cap [\beta]$.

Riformulando con cura: un mondo w è in $[K * (\alpha \wedge \beta)]$ sse $w \in S_{\alpha \wedge \beta}$ e $w \vDash \alpha \wedge \beta$. Allora $w \vDash \alpha$, e poiché w appartiene alla sfera $S_{\alpha \wedge \beta}$ che interseca $[\alpha]$ e S_α è la sfera minima che interseca $[\alpha]$, si ha che $w \in S_{\alpha \wedge \beta}$ con $S_\alpha \subseteq S_{\alpha \wedge \beta}$, dunque qualunque w in $S_\alpha \cap [\alpha] \cap [\beta]$ — che è $[(K * \alpha) + \beta]$ — è a fortiori in $S_{\alpha \wedge \beta} \cap [\alpha] \cap [\beta]$. La verifica di (K*8) è duale e segue dall'ipotesi che $\neg\beta \notin K * \alpha$, che semanticamente equivale a $S_\alpha \cap [\alpha] \not\subseteq [\neg\beta]$, ovvero $S_\alpha \cap [\alpha] \cap [\beta] \neq \varnothing$, da cui $S_{\alpha \wedge \beta} \subseteq S_\alpha$. ✓

La direzione **da postulati a sfere** — costruire esplicitamente un sistema di sfere a partire da un'operazione che soddisfa i postulati — è la parte costruttivamente più ricca. Grove definisce l'ordinamento di plausibilità su W ponendo:

$w_1 \leqslant_* w_2$ sse per ogni α: se $w_2 \in [K * \alpha]$ allora $w_1 \in [K * \alpha]$

In parole: w_1 è almeno tanto plausibile quanto w_2 se, ogni volta che w_2 viene selezionato dalla revisione per α, anche w_1 viene selezionato. Questo ordinamento è un preordine totale — la totalità discende da (K*2) e (K*4) — e il sistema di sfere associato è definito dai livelli di questa relazione. La verifica che l'operazione di revisione definita da questo sistema di sfere coincida con * originaria usa i postulati (K*3) e (K*4) per il caso non conflittuale, e (K*5) per garantire la consistenza delle selezioni.

Movimento IV — Entrenchment epistemico e ordinamento duale

La semantica delle sfere offre una rappresentazione *esogena* della struttura doxastica : l'ordinamento di plausibilità è definito sui mondi possibili, entità esterne al corpus dell'agente. Esiste però una rappresentazione *endogena*, che articola la struttura della revisione direttamente nei termini degli enunciati che l'agente crede: è la teoria degli *ordinamenti di entrenchment epistemico* sviluppata da Gärdenfors e Makinson nel 1988, in un lavoro complementare a quello di Grove.

L'intuizione di fondo è che non tutte le credenze di un agente abbiano lo stesso peso: alcune sono più centrali, più radicate, più difficili da abbandonare di altre. Quando un'informazione contraddittoria giunge, è le credenze *meno entrenched* — meno solide, più periferiche — che vengono sacrificate per prime. L'entrenchment è dunque una misura di quanto resistente alla revisione sia ogni singola credenza.

Formalmente, un *ordinamento di entrenchment epistemico* per K è una relazione binaria $\leq_e$ su L (l'insieme degli enunciati) che soddisfa le seguenti condizioni:

(EE1) $\leq_e$ è un preordine totale (riflessiva e transitiva, con $\alpha \leq_e \beta$ o $\beta \leq_e \alpha$ per ogni α, β)

(EE2) *Dominanza:* Se $\alpha \vdash \beta$, allora $\alpha \leq_e \beta$ (le conseguenze di un enunciato sono almeno tanto entrenched quanto l'enunciato stesso)

(EE3) *Congiuntività:* $\alpha \leq_e (\alpha \wedge \beta)$ o $\beta \leq_e (\alpha \wedge \beta)$ (una congiunzione è almeno tanto entrenched quanto il meno entrenched dei suoi congiunti)

(EE4) *Non trivialità:* Se K è consistente, allora per ogni $\alpha \in K$: $\alpha <_e \beta$ per qualche β (in un corpus consistente, non tutte le credenze hanno il massimo entrenchment)

(EE5) *Massimalità:* $\beta \leq_e \alpha$ per ogni β sse $\vdash \alpha$ (i teoremi logici — tautologie — sono i soli enunciati con entrenchment massimale assoluto)

Dove $\alpha <_e \beta$ significa $\alpha \leq_e \beta$ e non $\beta \leq_e \alpha$, ovvero α è *strettamente meno entrenched* di β.

Il legame tra l'ordinamento di entrenchment e l'operazione di revisione è espresso dal seguente *criterio C*:

(C) $\beta \in K * \alpha$ sse $\beta \in Cn(\varnothing)$ o $(\neg\alpha <_e (\neg\alpha \vee \beta))$

Il criterio C si legge così: un enunciato β appartiene al corpus revisionato $K * \alpha$ se e soltanto se β è una tautologia, oppure la negazione di α è *strettamente meno entrenched* della disgiunzione $\neg\alpha \vee \beta$. Questa seconda condizione ha un significato trasparente: per ritenere β dopo aver revisionato per α, si deve essere più disposti ad abbandonare $\neg\alpha$ che a cedere su β — ovvero, β deve essere più solido di $\neg\alpha$ agli occhi dell'agente.

Il *teorema di rappresentazione per l'entrenchment* afferma:

Teorema (Gärdenfors–Makinson, 1988). Un'operazione $*$ soddisfa (K*1)–(K*8) se e soltanto se esiste un ordinamento di entrenchment epistemico $\leq_e$ che soddisfa (EE1)–(EE5) e tale che (C) valga per ogni α, β.

La dimostrazione della direzione **da entrenchment a postulati** procede verificando i postulati uno per uno usando (C). Mostriamo (K*5) come caso esemplare: se $\nvdash \neg\alpha$, allora $\neg\alpha$ è consistente, e per (EE5) esiste β tale che $\neg\alpha <_e \beta$; in particolare, $\neg\alpha <_e \neg a \vee \neg a = \neg a$ *è falso, ma* $\neg a <_e (\neg a \vee \top) = \top$*, e usando (EE4) troviamo che a stesso — se* $a \in K$ *— ha un certo entrenchment positivo. La verifica formale che* $K * a \neq K\bot$ *usa il fatto che* $a \in K * a$ *per (K*2) e che se* $K * a = K\bot$ allora $\neg\alpha \in K * \alpha$, ma $\neg\alpha \in K * \alpha$ richiederebbe per (C) che $\neg\alpha \notin \mathrm{Cn}(\varnothing)$ (soddisfatto) e $\neg\alpha <_e (\neg\alpha \vee \neg\alpha) = \neg\alpha$, che è impossibile per riflessività di $\leq_e$. ✓

Per la direzione **da postulati a entrenchment**, Gärdenfors e Makinson costruiscono l'ordinamento di entrenchment *a partire* dall'operazione di revisione mediante la seguente definizione:

$\alpha \leq_e \beta$ sse $\alpha \notin K * \neg(\alpha \wedge \beta)$ o K è inconsistente

Ovvero: α è al massimo tanto entrenched quanto β se, nel revisionare per la negazione della loro congiunzione, si perde α (ma non necessariamente β). La perdita di α nella revisione per $\neg(\alpha \wedge \beta)$ indica che, quando si deve scegliere cosa abbandonare tra α e β, è α il sacrificato — dunque è β la credenza più radicata.

La verifica di (EE3) — la congiuntività — è la più delicata e richiede entrambi i postulati supplementari. Si deve mostrare che $\alpha \leq_e (\alpha \wedge \beta)$ o $\beta \leq_e (\alpha \wedge \beta)$, ovvero che almeno uno tra α e β è non più entrenched della loro congiunzione. Per definizione, $\alpha \leq_e (\alpha \wedge \beta)$ sse $\alpha \notin K * \neg(\alpha \wedge (\alpha \wedge \beta)) = K * (\neg\alpha \vee \neg\beta)$. Usando (K*7) e (K*8) e la proprietà di minima mutilazione congiunta, si mostra che almeno uno tra α e β deve essere assente dal corpus revisionato per $\neg\alpha \vee \neg\beta$, il che equivale a dire che almeno uno dei due ha un entrenchment non superiore alla congiunzione. ✓

Movimento V — Equivalenza delle tre rappresentazioni e geometria dell'epistemologia

---

A questo punto disponiamo di tre caratterizzazioni equivalenti della stessa struttura: i postulati AGM in forma assiomatica, le sfere di Grove in forma geometrica-topologica, e l'entrenchment di Gärdenfors–Makinson in forma ordinale-endogena. Questa triplice equivalenza non è un accidente tecnico: rivela che la razionalità della revisione doxastica ha una struttura profondamente geometrica.

La connessione tra le sfere di Grove e l'entrenchment epistemico è data dalla seguente corrispondenza: se $\leqslant$ è il preordine di plausibilità su W e $\leq_e$ è l'ordinamento di entrenchment su L, allora:

$\alpha \leq_e \beta$ *sse* $min_{\leqslant}([\neg\alpha]) \subseteq min_{\leqslant}([\neg(\alpha \wedge \beta)])$ oppure $min_{\leqslant}([\neg\beta]) \supseteq min_{\leqslant}([\neg(\alpha \wedge \beta)])$

dove $min_{\leqslant}(X) = \{w \in X \mid \forall v \in X: v \leqslant w \Rightarrow w \leqslant v\}$ è l'insieme dei mondi minimali in X rispetto a $\leqslant$. In termini più intuitivi: α è al massimo tanto entrenched quanto β se i mondi più plausibili che falsificano α sono al di fuori dei mondi più plausibili che falsificano $\alpha \wedge \beta$ — ovvero, α è più facile da falsificare nell'ordinamento di plausibilità.

Questa corrispondenza può essere espressa anche in termini di funzioni di selezione. Se $f : 2^W \to 2^W$ è la funzione che ad ogni insieme X di mondi associa i mondi minimali in X rispetto a $\leqslant$, allora:

$[K * \alpha] = f([\neg K] \cup [\alpha])$ quando K è consistente e α è compatibile con K $[K * \alpha] = f([\alpha])$ quando $\neg\alpha \in K$ (caso genuinamente revisionale)

dove $[\neg K] = W \setminus [K]$ è l'insieme dei mondi non compatibili con K. Il primo caso — quando K è compatibile con α — è ancora una volta l'identità tra revisione ed espansione. Il secondo caso — il cuore della teoria — mostra che revisionare per α equivale a selezionare i mondi più plausibili tra tutti quelli che rendono vera α.

La struttura di f è quella di una *selezione per minimalità*, e la condizione che essa soddisfi affinché la revisione risultante rispetti i postulati AGM è esattamente che il preordine $\leqslant$ sia *ben fondato* — ogni insieme non vuoto abbia elementi minimali — e *totale*. La ben fondatezza garantisce che $min_{\leqslant}([\alpha]) \neq \varnothing$ per ogni α consistente, evitando il caso patologico in cui non vi siano mondi selezionabili.

Vi è qui un'analogia profonda con la semantica condizionale di David Lewis per i condizionali controfattuali. Lewis aveva proposto nel 1973 di interpretare "se fosse α, sarebbe β" ($\alpha \mathrel{\square\!\!\rightarrow} \beta$) come vera in un mondo w se e soltanto se tutti i mondi α-più-vicini a w rendono vero β. La struttura delle sfere che Lewis usa per definire la prossimità è formalmente identica a quella di Grove: un sistema di sfere centrato su w, totalmente ordinato, con w come elemento minimo. La differenza è che nella semantica di Lewis il centro varia di mondo in mondo — ogni mondo ha il suo sistema di sfere — mentre nella semantica di Grove il centro è fisso (è il corpus K dell'agente) e le sfere rappresentano la struttura doxastica di un singolo soggetto.

Questa consonanza non è superficiale. Suggerisce che la razionalità della revisione doxastica e la semantica dei condizionali appartengono alla stessa famiglia di strutture formali: entrambe sono istanze della geometria degli spazi di plausibilità ordinati. Gärdenfors aveva già notato — in un argomento poi diventato controverso — che questa connessione implica una relazione tra la revisione per α e la credenza nel condizionale "se α fosse vero, allora β": specificamente, $\beta \in K * \alpha$ dovrebbe equivalere a $(\alpha \mathrel{\square\!\!\rightarrow} \beta) \in K$. Questo *teorema di Ramsey*, come viene chiamato, è stato

oggetto di una critica devastante di Gärdenfors stesso, che ha dimostrato — in quello che è noto come il *teorema di impossibilità di Gärdenfors* — che nessuna operazione di revisione che soddisfi i postulati AGM può rendere simultaneamente soddisfacente questa equivalenza in modo non triviale. Il teorema di impossibilità è uno dei risultati filosoficamente più profondi dell'epistemologia formale contemporanea, e merita trattazione separata.

Movimento VI — Il teorema di impossibilità di Gärdenfors

Poche dimostrazioni nell'epistemologia formale hanno la struttura di un *reductio* tanto preciso quanto quello che Gärdenfors pubblicò nel 1986, due anni prima del completamento del quadro semantico di Grove. L'obiettivo era verificare se la connessione intuitiva tra revisione doxastica e condizionali sogguntivi potesse essere formalizzata in modo coerente. Il risultato fu negativo, e la negatività del risultato è filosoficamente istruttiva quanto lo sarebbe stata una risposta affermativa.

Il punto di partenza è il cosiddetto *test di Ramsey*, così denominato perché Frank Ramsey aveva suggerito, in un frammento postumo del 1929, che il modo più naturale di valutare un condizionale "se α, allora β" consiste nell'ipotizzare α, aggiornare le proprie credenze alla luce di questa ipotesi, e verificare se β risulti credibile nel corpus così modificato. La proposta di Ramsey non era formalmente sviluppata, ma la sua struttura è quella di un collegamento sistematico tra la relazione di revisione $*$ e la relazione di credenza condizionale.

Formalmente, il collegamento viene espresso attraverso quello che Gärdenfors chiama il *principio di Ramsey* (PR):

(PR) $(\alpha \rightarrow \beta) \in K$ sse $\beta \in K * \alpha$

dove $\rightarrow$ denota il condizionale sogguntivo del linguaggio (che si può intendere come il $\Box\!\!\rightarrow$ lewisiano o, più debolmente, come un operatore condizionale non classico da specificare). Il principio dice: un agente crede al condizionale "se α allora β" se e soltanto se, dopo aver revisionato le proprie credenze supponendo α vera, crede β.

L'eleganza del principio è evidente: fornisce una semantica doxastica dei condizionali che li ancora all'operazione cognitiva di revisione piuttosto che a valutazioni di mondi possibili esterne all'agente. È una proposta *epistemicamente internalista*, nel senso che l'intera analisi si svolge all'interno dello spazio delle credenze dell'agente.

Gärdenfors dimostrò che (PR), congiunto con i postulati AGM e con un principio minimale di chiusura per i condizionali, porta a una contraddizione. La dimostrazione è la seguente.

Si supponga che K sia un corpus consistente che non contenga né α né $\neg\alpha$ né β né $\neg\beta$, con α, β, $\neg\alpha$, $\neg\beta$ tutti consistenti con K. Per (PR) e per il postulato di vacuità

(K*4) — che identifica la revisione con l'espansione quando non vi è conflitto — si ha:

$(\alpha \rightarrow \beta) \in K$ sse $\beta \in K + \alpha = Cn(K \cup \{\alpha\})$

Ma $\beta \notin Cn(K \cup \{\alpha\})$ per ipotesi (essendo β e α indipendenti da K), dunque $(\alpha \rightarrow \beta) \notin K$. Analogamente $(\alpha \rightarrow \neg\beta) \notin K$. Sin qui nessun problema.

Si consideri ora il corpus $K' = K + (\alpha \rightarrow \beta)$, ovvero K espanso con la credenza nel condizionale. Per (PR) applicato a K':

$\beta \in K' * \alpha$

Ora, K' contiene $(\alpha \rightarrow \beta)$ per costruzione. Si applichi di nuovo (PR) per passare da K' a $K' * (\alpha \wedge \neg\beta)$. Per i postulati supplementari (K*7) e (K*8), e poiché $\neg\beta$ è compatibile con $K' * \alpha$ (si verifica che $K' * \alpha$ non contenga β in modo necessario — questo è il punto cruciale che richiede un'ipotesi aggiuntiva di *robustezza* della revisione), si ottiene:

$K' * (\alpha \wedge \neg\beta) = (K' * \alpha) + \neg\beta$

E poiché $\beta \in K' * \alpha$ per quanto mostrato sopra, si avrebbe $\beta \in (K' * \alpha) + \neg\beta$, il che produce immediatamente $\bot \in (K' * \alpha) + \neg\beta$ — dunque il corpus $(K' * \alpha) + \neg\beta$ è inconsistente. Ma $(K' * \alpha) + \neg\beta = K' * (\alpha \wedge \neg\beta)$, e per il postulato (K*5) la revisione per $\alpha \wedge \neg\beta$ deve essere consistente (essendo $\alpha \wedge \neg\beta$ consistente). Contraddizione. ✗

La versione più raffinata della dimostrazione, quella che Gärdenfors presenta nell'articolo del 1986 e che ha il massimo rigore, assume una formulazione leggermente diversa delle ipotesi, ma la struttura logica è identica: si usa (PR) per dedurre che il corpus revisionato deve contenere sia β sia $\neg\beta$, il che viola (K*5).

La diagnosi filosofica del teorema è dibattuta. Una prima lettura — quella di Gärdenfors stesso — è che i condizionali sogguntivi non possono essere trattati come credenze ordinarie all'interno di un corpus doxastico che soddisfi i postulati AGM: i condizionali sfuggono alla logica della revisione perché il loro contenuto dipende strutturalmente dall'operazione di revisione stessa, creando una circolarità che il sistema non può reggere. Una seconda lettura — difesa da autori come Levi — è che il principio di Ramsey è semplicemente troppo forte: ciò che esso esprime non è l'analisi della credenza condizionale ordinaria, bensì una forma di *scommessa condizionale* o *impegno ipotetico* che non ha lo stesso statuto logico delle credenze nel corpus. Una terza lettura — più radicale — è che il teorema rivela un limite intrinseco della rappresentazione doxastica classica: la nozione di corpus come teoria chiusa è troppo statica per catturare la struttura degli stati mentali che intervengono nell'inferenza condizionale.

Movimento VII — Il problema della revisione iterata

Tra i limiti più discussi della teoria AGM vi è quello che Darwiche e Pearl hanno denominato il *problema della revisione iterata*. Il punto è concettualmente semplice ma formalmente grave: i postulati AGM specificano come si passa da K a K * α, ma non dicono nulla su come debba essere strutturata l'operazione di revisione *successiva* (K * α) * β. Il problema non è meramente tecnico: riguarda la *dinamica* dell'identità epistemica di un agente che riceve informazioni in sequenza.

Si consideri un agente che crede inizialmente che un dado a sei facce mostri un numero pari (credenza P), e che riceva dapprima la notizia che il numero è maggiore di tre (revisione per α: "il numero è 4, 5 o 6"), e poi la notizia che il numero è dispari (revisione per β: "il numero è 1, 3 o 5"). La prima revisione produce un corpus K * α compatibile con {4, 5, 6}, e la seconda revisione dovrebbe produrre (K * α) * β compatibile con {5}. Questo è il risultato intuitivamente corretto.

Ma si consideri un agente alternativo che parte dallo stesso corpus K, non riceve la prima notizia, e revisiona direttamente per α ∧ β — ovvero, viene informato che il dado mostra un numero maggiore di tre *e* dispari, dunque mostra 5. Anche questa revisione dovrebbe produrre un corpus compatibile con {5}. I due processi — revisione sequenziale e revisione congiunta — dovrebbero convergere, almeno in questo caso.

I postulati AGM non garantiscono questa convergenza in generale, perché non specificano quale struttura di plausibilità il corpus revisionato K * α erediti dalla struttura di plausibilità originaria di K. In termini di sfere di Grove: quando si revisiona K per α, il sistema di sfere S si restringe e si aggiorna in S', ma i postulati AGM non determinano univocamente S' a partire da S. Diverse scelte di S' — tutte compatibili con i postulati — producono diversi risultati per la revisione successiva (K * α) * β.

Darwiche e Pearl formalizzarono questo problema nel 1997 proponendo quattro postulati aggiuntivi, designati (C1)–(C4), che governano il comportamento della revisione iterata. Usando la notazione K_α = K * α per compattezza, i postulati sono:

(C1) Se β ∈ K_α * α, allora β ∈ (K_α) * β...

Riformuliamo con la notazione standard. Sia ψ_K il *condizionale doxastico* di K, ovvero la funzione che ad ogni formula α associa il corpus K * α. Darwiche e Pearl lavorano direttamente a livello degli ordinamenti di plausibilità: dato un ordinamento $\leqslant_K$ *e una formula a, l'ordinamento revisionato* $\leqslant_{K*\alpha}$ deve soddisfare:

(C1) Se $w_1 \models \alpha$ e $w_2 \models \alpha$, allora $w_1 \leqslant_{K*\alpha} w_2$ sse $w_1 \leqslant_K w_2$

(C2) Se $w_1 \not\models \alpha$ e $w_2 \not\models \alpha$, allora $w_1 \leqslant_{K*\alpha} w_2$ sse $w_1 \leqslant_K w_2$

(C3) Se $w_1 \models \alpha$ e $w_2 \not\models \alpha$, allora $w_1 <_K w_2 \Rightarrow w_1 <_{K*\alpha} w_2$

(C4) Se $w_1 \models \alpha$ e $w_2 \not\models \alpha$, allora $w_1 \leqslant_{K*\alpha} w_2 \Longrightarrow w_1 \leqslant_K w_2$ o $w_1 \sim_K w_2$

dove $w_1 \sim_K w_2$ indica che w_1 e w_2 sono equivalenti nell'ordinamento originario ($w_1 \leqslant_K w_2$ e $w_2 \leqslant_K w_1$).

Il significato di questi quattro vincoli è perspicuo. (C1) afferma che la revisione per α non altera l'ordinamento relativo tra i mondi che già soddisfano α: l'ordine interno ai mondi α-compatibili è preservato. (C2) afferma la medesima cosa per i mondi α-incompatibili: anche l'ordine tra i mondi che falsificano α è preservato dalla revisione. (C3) afferma che se un mondo α-compatibile era più plausibile di un mondo α-incompatibile prima della revisione, continua ad esserlo dopo — l'elevazione dell'intera classe α alla primazia non fa che rafforzare una precedenza già esistente. (C4) afferma che se dopo la revisione un mondo α-compatibile è più plausibile di uno α-incompatibile, allora o lo era già prima, oppure erano equivalenti prima: la revisione non inverte mai la precedenza dei mondi α-incompatibili sui compatibili a meno che non vi sia indifferenza iniziale.

L'eleganza di (C1)–(C4) consiste nel fatto che essi descrivono la revisione come una *ripartizione e riordinamento* dell'ordinamento originario: si separa il piano dei mondi α-compatibili da quello degli α-incompatibili, si porta il primo al di sopra del secondo, ma si preservano gli ordinamenti interni a ciascun piano. Questa è esattamente la struttura di ciò che Darwiche e Pearl chiamano *condizionamento*, in analogia con il condizionamento probabilistico, e che la letteratura successiva ha adottato sotto il nome di *revisione naturale* o *revisione lessicografica* a seconda delle varianti.

La versione *lessicografica* della revisione — talvolta attribuita a Nayak — è la più conservativa: essa colloca tutti i mondi α-compatibili prima di tutti gli α-incompatibili, preservando gli ordinamenti interni, ed è caratterizzata dall'ulteriore postulato:

(Lex) Se $w_1 \models \alpha$ e $w_2 \not\models \alpha$, allora $w_1 <_{K*\alpha} w_2$

dove $<$ denota la precedenza stretta. Questo è più forte di (C3) e (C4) congiuntamente, perché stabilisce una precedenza assoluta e incondizionata dei mondi α-compatibili, indipendentemente da quanto fossero implausibili nella struttura originaria. Il prezzo di questa semplicità è una forma di *rigidità epistemica*: il corpus revisionato tratta la nuova informazione come inoppugnabile, indipendentemente da quanto possa sembrare sorprendente.

La revisione *naturale* di Boutilier è invece più moderata: essa distingue tra il caso in cui α conferma le aspettative dell'agente (ovvero i mondi α-compatibili più plausibili sono già tra i più plausibili in assoluto) e il caso in cui α le sorprende. Solo nel secondo caso la struttura di plausibilità viene ristrutturata significativamente. La revisione naturale soddisfa (C1)–(C4) ma non (Lex), ed è caratterizzata dalla condizione:

(Nat) $\leqslant_{K*\alpha}$ è il preordine più coarse-grained (più grossolano) compatibile con (C1)–(C4)

Nel senso che colloca i mondi α-compatibili al di sopra di quelli α-incompatibili solo nella misura strettamente necessaria a soddisfare i vincoli di Darwiche e Pearl, senza ulteriori ristrutturazioni.

Movimento VIII — L'assiomatica di Darwiche–Pearl e il teorema di rappresentazione iterato

---

Il contributo tecnico più rilevante di Darwiche e Pearl non è soltanto l'identificazione di (C1)–(C4) come condizioni naturali, ma la dimostrazione che esse sono categoricamente distinte dai postulati AGM — ovvero, non sono deducibili da questi — e che, congiunte con i postulati AGM, hanno un *teorema di rappresentazione* preciso.

La versione assiomatica di (C1)–(C4) a livello delle operazioni di revisione — piuttosto che degli ordinamenti sottostanti — è la seguente. Si usa la notazione $K_{\alpha} = K * \alpha$ e $(K_{\alpha})_{\beta} = (K * \alpha) * \beta$:

(C1′) Se $\beta \vdash \alpha$, allora $(K_{\alpha})_{\beta} = K_{\beta}$

(C2′) Se $\beta \vdash \neg\alpha$, allora $(K_{\alpha})_{\beta} = K_{\beta}$

(C3′) Se $\alpha \in K_{\beta}$, allora $\alpha \in (K_{\alpha})_{\beta}$

(C4′) Se $\neg\alpha \notin K_{\beta}$, allora $\neg\alpha \notin (K_{\alpha})_{\beta}$

Il passaggio dalle versioni sugli ordinamenti (C1)–(C4) a quelle sulle operazioni (C1′)–(C4′) richiede una verifica non banale, che Darwiche e Pearl svolgono nel dettaglio. Il caso più illuminante è (C1′): se β implica α, allora tutti i mondi β-compatibili sono anche α-compatibili. Revisionare per α e poi per β, nella lettura di (C1), deve preservare l'ordinamento tra i mondi β-compatibili (per (C1)) e tra quelli β-incompatibili (per (C2)), ma tutti i mondi che precedono variano entro le classi α-compatibili e α-incompatibili. Poiché $\beta \vdash \alpha$, non vi è mai conflitto tra le due revisioni nel senso di mondi che soddisfano β ma non α. La revisione per α, in questo caso, non disturba nulla di ciò che sarà rilevante per la successiva revisione per β, dunque il risultato finale è lo stesso che si otterrebbe revisionando direttamente per β a partire da K.

Il *teorema di rappresentazione DP* afferma:

Teorema (Darwiche–Pearl, 1997). Un'operazione di revisione iterata $*$ soddisfa (K*1)–(K*8) e (C1′)–(C4′) se e soltanto se esiste una famiglia $\{\preccurlyeq_K\}_K$ di preordini totali e ben fondati — uno per ogni corpus K — tale che:

(i) $[K] = \min_{\preccurlyeq_K}(W)$ (i mondi più plausibili secondo $\preccurlyeq_K$ sono esattamente quelli compatibili con K)

(ii) [K * α] = min_{⩽_K}([α]) (la revisione seleziona i mondi α-compatibili più plausibili)

(iii) ⩽_{K*α} soddisfa (C1)–(C4) rispetto a ⩽_K

La condizione (iii) è quella che aggiunge contenuto rispetto al teorema di Grove: non soltanto ogni corpus K è associato a un ordinamento di plausibilità, ma questi ordinamenti sono *mutuamente coerenti* nel senso specificato da (C1)–(C4). La famiglia degli ordinamenti forma così una struttura coesa — una *politica epistemica* dell'agente — piuttosto che una raccolta di strutture indipendenti.

Una conseguenza importante è che i postulati DP non determinano univocamente ⩽_{K*α} a partire da ⩽_K: esiste un'intera famiglia di ordinamenti revisionati compatibili con (C1)–(C4), e la scelta tra di essi rappresenta diversi stili di aggiornamento epistemico. Questa indeterminazione è filosoficamente significativa: suggerisce che la razionalità della revisione iterata lascia spazio a legittime differenze di stile cognitivo, non riducibili a criteri puramente logici.

Il problema della revisione iterata mette in luce una tensione più profonda nella teoria AGM. I postulati originali trattano il corpus K come un'entità statica, un punto nello spazio delle teorie: essi specificano come si passa da un punto all'altro, ma non forniscono all'agente una struttura che sopravviva alla transizione. I postulati DP, aggiungendo vincoli sull'evoluzione degli ordinamenti di plausibilità, cominciano a trattare l'agente come una *traiettoria* nello spazio epistemico piuttosto che come un punto. Ma anche questa prospettiva è incompleta: non dice nulla sulla dipendenza della struttura di plausibilità dalla *storia* delle revisioni passate, né su come l'agente debba comportarsi di fronte a sequenze di informazioni potenzialmente contraddittorie che si annullano a vicenda.

Movimento IX — Il condizionamento bayesiano come caso limite

La teoria AGM si sviluppa interamente nel quadro della logica classica e tratta la credenza come una relazione binaria: un agente crede una proposizione oppure non la crede, senza gradi intermedi. Questa scelta, consapevolmente idealizzante, ha permesso una formalizzazione rigorosa e matematicamente ricca. Ma l'epistemologia ha anche una tradizione alternativa, altrettanto rigorosa, che tratta le credenze come entità graduali: la teoria della probabilità soggettiva, nella versione che va da Ramsey e de Finetti fino al bayesianismo contemporaneo. La domanda se queste due tradizioni siano rivali o complementari è una delle questioni aperte più fertili dell'epistemologia formale.

Il nesso formale tra le due prospettive è stato studiato da numerosi autori, e il risultato più chiaro è che il condizionamento bayesiano — l'aggiornamento della funzione di probabilità soggettiva per regola di Bayes — si comporta, in un senso preciso, come

un caso limite della revisione AGM quando le probabilità convergono verso zero e uno.

Si consideri uno spazio di probabilità (W, $\mathscr{A}$, P), dove W è l'insieme dei mondi possibili, $\mathscr{A}$ è una σ-algebra su W, e P è una funzione di probabilità soggettiva. L'aggiornamento bayesiano per la proposizione α (con P(α) > 0) produce la probabilità condizionale:

$P(\cdot \mid \alpha) : \beta \mapsto P(\beta \cap \alpha) / P(\alpha)$

Questa regola di condizionamento è ben definita ed è l'unico modo di aggiornare che soddisfi simultaneamente la coerenza dinamica (Dutch book diacronico) e la regola della catena. La sua struttura è quella di una *riscalatura*: si azzera la probabilità di tutti i mondi α-incompatibili e si riscalano proporzionalmente le probabilità dei mondi α-compatibili affinché la misura totale rimanga unitaria.

Ora si consideri la famiglia di distribuzioni di probabilità P_ε, parametrizzata da $\varepsilon > 0$, definita a partire da un preordine di plausibilità $\preccurlyeq$ su W mediante:

$P_\varepsilon(w) = \varepsilon^{\mathrm{rank}(w)} \cdot Z_\varepsilon^{-1}$

dove rank(w) è il rango di w nell'ordinamento $\preccurlyeq$ (il minimo rango è 0, per i mondi più plausibili) e $Z_\varepsilon = \Sigma_{w \in W} \varepsilon^{\mathrm{rank}(w)}$ è la costante di normalizzazione. Questa famiglia è nota come distribuzione di *misura di possibilità esponenziale* o *distribuzione epsilon-ranking*, ed è centrale nel lavoro di Adam Grove, Halpern, e Spohn.

Il comportamento limite è il seguente. Per $\varepsilon \to 0$, la massa di probabilità si concentra sui mondi di rango minimo:

$\lim_{\varepsilon \to 0} P_\varepsilon(\alpha) = 1$ se e soltanto se $\min_{\preccurlyeq}(W) \subseteq [\alpha]$

ovvero P_ε attribuisce probabilità tendente a 1 a α se e soltanto se tutti i mondi più plausibili rendono vera α. Questo è esattamente il criterio di credenza nell'ordinamento di plausibilità: $K = \{\alpha \mid \min_{\preccurlyeq}(W) \subseteq [\alpha]\}$.

Per quanto riguarda l'aggiornamento, il condizionamento bayesiano di P_ε per α produce:

$P_\varepsilon(\beta \mid \alpha) = \Sigma_{w \in [\beta] \cap [\alpha]} \varepsilon^{\mathrm{rank}(w)} / \Sigma_{w \in [\alpha]} \varepsilon^{\mathrm{rank}(w)}$

Nel limite $\varepsilon \to 0$, questo espressione converge a:

$\lim_{\varepsilon \to 0} P_\varepsilon(\beta \mid \alpha) = 1$ sse $\min_{\preccurlyeq}([\alpha]) \subseteq [\beta]$

che è esattamente la condizione di credenza nella revisione AGM: $\beta \in K * \alpha$ sse i mondi α-più-plausibili soddisfano tutti β. Il condizionamento bayesiano, applicato alla famiglia P_ε e portato al limite, replica precisamente la revisione per ordinamento di plausibilità.

Questo risultato — che si trova in forma esplicita in Grove (1988) e poi generalizzato da Halpern (2003) — ha un'interpretazione filosofica duale. Da un lato, suggerisce che la revisione AGM è la versione *qualitativa* del condizionamento bayesiano: si ottiene da esso eliminando l'informazione quantitativa sulle probabilità e ritenendo soltanto l'ordine relativo di plausibilità. Dall'altro lato, suggerisce che il condizionamento bayesiano è un *arricchimento* della revisione AGM: aggiunge informazione quantitativa all'ordine qualitativo, permettendo di distinguere non soltanto quali mondi siano più plausibili ma quanto lo siano.

La questione normativa — quale delle due rappresentazioni sia più adeguata per un agente razionale — non ha risposta univoca. Gli argomenti a favore della rappresentazione probabilistica sono noti: coerenza diacronica garantita dai teoremi di Dutch book, compatibilità con la teoria della decisione, possibilità di aggiornamento parziale (in cui la nuova informazione sposta le probabilità ma non le porta ai valori estremi). Gli argomenti a favore della rappresentazione ordinale sono altrettanto cogenti: minore onere computazionale, maggiore generalità (ogni ordinamento di plausibilità corrisponde a una classe di distribuzioni probabilistiche, non a una sola), maggiore robustezza rispetto all'assegnazione numerica delle probabilità, che in molti contesti reali è arbitraria o impossibile.

La proposta di Spohn di usare le *funzioni di ranking* — funzioni $\kappa : W \to \mathbb{N} \cup \{\infty\}$ che assegnano a ogni mondo un rango naturale (con $\kappa(w) = 0$ per i mondi più plausibili) — cerca di occupare un posto intermedio: è più ricca degli ordinamenti puri (perché permette di distinguere quanto sia improbabile un mondo in termini cardinali) ma più povera della probabilità completa (perché usa i numeri naturali anziché il reale [0,1]).

La regola di aggiornamento di Spohn è:

$\kappa_{*\alpha}(w) = \kappa(w) - \min_{v \in [\alpha]} \kappa(v)$ per $w \in [\alpha]$ $\kappa_{*\alpha}(w) = \kappa(w) + c_\alpha$ per $w \notin [\alpha]$

dove c_α è una costante positiva che esprime il "peso" con cui viene ricevuta l'informazione α. Quando $c_\alpha \to \infty$, questa regola converge alla revisione AGM (tutti i mondi α-incompatibili vengono rimossi del tutto); per valori finiti di c_α, produce aggiornamenti più graduali che permettono di "ricordare" anche i mondi scartati e potenzialmente reintegrarli in revisioni future.

Movimento X — Coerenza, fondazionalismo e il problema del radicamento

---

La teoria AGM, nelle sue diverse rappresentazioni, è una teoria della *coerenza* della revisione doxastica : specifica come un corpus di credenze debba cambiare per restare internamente consistente di fronte a nuove informazioni. Ma la domanda sulla coerenza non esaurisce le questioni normative dell'epistemologia. Accanto a essa si

pone il problema del *radicamento* — in inglese *grounding* — che riguarda non la coerenza interna del corpus ma il suo *ancoraggio* alla realtà.

Un corpus K può essere internamente coerente, soddisfare tutti i postulati AGM, e tuttavia contenere unicamente credenze false o del tutto scollegate dall'evidenza empirica. I postulati AGM sono indifferenti al contenuto delle credenze: si applicano ugualmente a un corpus di credenze vere e a uno di credenze false, purché entrambi siano internamente consistenti. Questa neutralità rispetto al contenuto è una virtù formale — permette alla teoria di avere massima generalità — ma è anche un limite filosofico: una teoria della razionalità epistemica che non discrimina tra corpora ben ancorati all'evidenza e corpora del tutto arbitrari è, da un punto di vista esternalista, incompleta.

Il problema del radicamento si presenta in forma acuta quando si considerano gli *input* della revisione. I postulati AGM trattano il nuovo dato α come un semplice enunciato che viene accolto nel corpus senza che la sua origine o affidabilità vengano rappresentate. Nella pratica cognitiva reale, tuttavia, la revisione è mediata da processi percettivi, testimonianze altrui, inferenze strumentali — ciascuno dei quali ha un proprio grado di affidabilità. Un agente razionale dovrebbe, in linea di principio, revisionare le proprie credenze in modo diverso a seconda di quanto affidabile sia la fonte dell'informazione ricevuta.

Questa critica — che ha i suoi portavoce più articolati in Isaac Levi e, da una prospettiva diversa, in Timothy Williamson — non è devastante per la teoria AGM, ma ne circoscrive il dominio di applicabilità. La risposta più naturale è che la teoria AGM descrive la struttura *condizionale* della revisione: dato che un agente accetta α come input, come deve aggiornare il proprio corpus? La domanda su *quando* debba accettare α come input appartiene a una teoria dell'evidenza distinta, che la teoria AGM non pretende di fornire.

Levi ha però una critica più specifica, che riguarda il trattamento della contrazione. Per Levi, la contrazione di una credenza da un corpus — il rinunciare volontariamente a una credenza che si possiede — è un'operazione fondamentalmente diversa dalla revisione, e i postulati AGM per la contrazione (derivati dall'identità di Harper) non catturano adeguatamente le condizioni di razionalità di questa operazione. In particolare, l'identità di Harper:

$$K - \alpha = K \cap (K * \neg\alpha)$$

implica che la contrazione sia determinata dalla revisione per la negazione — il che presuppone che l'agente sappia già come revisionare per $\neg\alpha$, e dunque che abbia già in mente una struttura di plausibilità completa. Ma secondo Levi, la contrazione pura — rinunciare a α senza ricevere alcuna informazione contraria — non richiede né presuppone alcuna propensione alla revisione: è un atto deliberativo autonomo che merita una teoria propria.

La proposta di Levi è quella di trattare la contrazione come la scelta di passare da K a un corpus $K - \alpha$ che sia un *sottoinsieme* di K massimale tra quelli che non implicano α.

La struttura di questa massimalità produce, quando K è finito, un'operazione ben definita — ma quando K è infinito (come nel caso dei corpora chiusi per conseguenza) si apre la questione della scelta tra i diversi sottoinsiemi massimali, nota come il problema delle *contrazioni partial meet*.

La *contrazione partial meet*, proposta da Alchourrón, Gärdenfors e Makinson nello stesso articolo del 1985, introduce una *funzione di selezione* γ che, ad ogni corpus K e ad ogni formula α, seleziona una sottofamiglia $\gamma(K, \alpha)$ dell'insieme $K \perp \alpha$ di tutti i sottoinsiemi massimali di K che non implicano α:

$K \perp \alpha = \{X \subseteq K \mid X \text{ non implica } \alpha, \text{ e per ogni } Y \text{ con } X \subset Y \subseteq K: Y \text{ implica } \alpha\}$

La contrazione partial meet è definita come:

$K -_\gamma \alpha = \bigcap \gamma(K, \alpha)$

ovvero l'intersezione di tutti i sottoinsiemi massimali selezionati da γ. Se γ seleziona tutti gli elementi di $K \perp \alpha$, si ottiene la *contrazione full meet* $K -_\cap \alpha = \bigcap(K \perp \alpha)$, che è la più conservativa ma anche la più debole (rimuove troppo). Se γ seleziona un unico elemento, si ottiene la *maxichoice contraction*, che è la più informativa ma viola i postulati di recupero in casi patologici.

La condizione che la funzione di selezione γ debba soddisfare perché la contrazione partial meet rispetti i postulati AGM per la contrazione è che essa sia *basata sull'entrenchment*: ovvero, che γ selezioni i sottoinsiemi di $K \perp \alpha$ che contengono gli enunciati più entrenched. Formalmente:

$X \in \gamma(K, \alpha)$ sse per ogni $Y \in K \perp \alpha$ e per ogni $\beta \in Y \setminus X$: $\beta <_e \delta$ per qualche $\delta \in X \setminus Y$

Questo criterio assicura che, tra tutte le possibili contrazioni, si scelga quella che preserva le credenze più solide — il che è esattamente l'intuizione originaria dell'entrenchment applicata all'operazione inversa della revisione.

Movimento XI — Riflessioni conclusive: razionalità, geometria e i limiti del formalismo

---

L'impresa intellettuale che abbiamo percorso in questi movimenti ha prodotto una struttura formale di notevole eleganza: i postulati AGM, le sfere di Grove, l'entrenchment di Gärdenfors–Makinson e i postulati di Darwiche–Pearl si compongono in un quadro coeso in cui la razionalità della revisione doxastica ha una caratterizzazione sia assiomatica sia geometrica. La triplice equivalenza tra le rappresentazioni non è una mera curiosità tecnica: rivela che la struttura della credenza razionale è, in un senso preciso, una struttura di *prossimità ordinata* — una metrica qualitativa dello spazio dei mondi possibili.

Questa geometria della credenza ha risvolti filosofici che vanno oltre l'epistemologia formale. Primo: essa suggerisce che la razionalità epistemica non è una proprietà monolitica ma dipende dalla struttura dell'intero spazio di plausibilità dell'agente — un agente non può essere valutato razionale o irrazionale sulla base di singole credenze isolate, ma soltanto sulla base della coerenza della sua struttura ordinale globale. Secondo: la geometria delle sfere rivela che la razionalità ha una topologia — la nozione di "credenza più vicina alla realtà" è una nozione topologica, non aritmetica. Terzo: il teorema di impossibilità di Gärdenfors mostra che questa geometria ha *curvature negative* rispetto ai condizionali: la struttura degli spazi di plausibilità non è compatibile con una semantica condizionale che sia simultaneamente intuitiva e coerente con i postulati di revisione.

Rimangono, al termine di questo percorso, alcune questioni aperte di prima grandezza. La prima riguarda la *revisione non-monotonica*: i postulati AGM presuppongono che l'agente debba accogliere l'input α incondizionatamente, ma nella pratica epistemica reale un agente può avere buone ragioni per *resistere* a un'informazione che contraddice credenze molto solidamente radicate. Una teoria della revisione che tenga conto della *resilienza* doxastica — la capacità di alcune credenze di sopravvivere a pressioni contrarie forti — richiederebbe modifiche sostanziali ai postulati.

La seconda questione aperta riguarda la *multi-agentezza*: i postulati AGM descrivono la revisione di un singolo agente isolato, ma la maggior parte della cognizione umana è cognitivamente distribuita — le credenze si formano attraverso interazioni sociali, testimonianze, dibattiti, autorità epistemiche. Una teoria della revisione doxastica che incorpori la dimensione sociale richiederebbe di trattare non soltanto la struttura di plausibilità di un singolo agente ma le relazioni tra le strutture di plausibilità di agenti diversi.

La terza questione, forse la più profonda, riguarda il rapporto tra la razionalità *sincronica* — la coerenza interna del corpus in un dato istante — e la razionalità *diacronica* — la coerenza dell'evoluzione del corpus nel tempo. I postulati AGM e DP affrontano aspetti di entrambe, ma non forniscono una teoria unificata. Il legame tra le due forme di razionalità, e le condizioni in cui possono entrare in conflitto, rimane uno dei problemi irrisolti più stimolanti dell'epistemologia formale contemporanea. La geometria della credenza che abbiamo costruito in questi movimenti è, in definitiva, la geometria di un agente che pensa in istanti successivi — e la domanda su come questi istanti si colleghino in una traiettoria razionalmente coerente è, insieme, una domanda logica, filosofica e, in ultima analisi, profondamente umana.

Movimento XII — Ragionamento non-monotono e l'eredità ordinale della revisione

---

La logica classica è monotona: se $\Gamma \vdash \alpha$, allora $\Gamma \cup \{\beta\} \vdash \alpha$ per qualsiasi β. Questa proprietà, ovvia dal punto di vista della deduzione formale, è cognitivamente irrealistica. Un medico che inferisce, dai sintomi di un paziente, la diagnosi di influenza, non mantiene necessariamente quella diagnosi quando viene aggiunta

l'informazione che il paziente è appena rientrato da una zona malarica. L'inferenza si *ritira* di fronte a nuova evidenza — ed è esattamente questa ritirabilità che la logica monotona non può catturare.

Il ragionamento non-monotono non è, dunque, una patologia logica da correggere, bensì una caratteristica strutturale del pensiero ordinario sotto incertezza. La domanda è come formalizzarlo senza perdere il rigore che la logica classica garantisce. Questa domanda ha generato, a partire dagli anni Ottanta, una famiglia di proposte — dalla default logic di Reiter alla circumscription di McCarthy, dalla Autoepistemic Logic di Moore fino al framework di Kraus, Lehmann e Magidor — che si sono progressivamente raffinate fino a identificare le condizioni minimali di razionalità del ragionamento non-monotono.

Il collegamento con la teoria AGM è strutturalmente profondo, e non si tratta di una semplice analogia. Entrambe le tradizioni — la teoria della revisione e la teoria del ragionamento non-monotono — sono state sviluppate in isolamento relativo per quasi un decennio, e soltanto a partire dal 1990 autori come Makinson, Gärdenfors e Lehmann hanno sistematicamente mostrato che esse sono, in un senso preciso, due facce della stessa struttura matematica sottostante: l'ordinamento di plausibilità sui mondi possibili.

La connessione si manifesta attraverso la nozione di *relazione di conseguenza non-monotona*. Una tale relazione, che si scrive $\alpha \mid\sim \beta$ (si legge "tipicamente, se α allora β" oppure "α inferisce tentativamente β"), è intesa come una generalizzazione del connettivo condizionale che cattura le inferenze che valgono *di regola* ma che possono essere sospese in presenza di informazioni eccezionali. La struttura formale di $\mid\sim$ è quella di una relazione binaria sull'insieme degli enunciati — o equivalentemente, una funzione $C : 2^L \to 2^L$ che ad ogni insieme di premesse associa l'insieme delle sue conseguenze non-monotone — e la domanda è quali postulati debba soddisfare perché possa dirsi *razionale*.

Movimento XIII — Il framework KLM: postulati e struttura preferenziale

---

Nel 1990, Sarit Kraus, Daniel Lehmann e Menachem Magidor pubblicarono in *Artificial Intelligence* un articolo che ha ridefinito i termini della discussione: "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics". Il contributo centrale è l'identificazione di una famiglia di postulati per $\mid\sim$ — i *postulati KLM* — e la dimostrazione che essi sono esattamente caratterizzati da una classe di strutture semantiche dette *modelli preferenziali*.

I postulati KLM per una relazione $\mid\sim$ sono i seguenti. Si usa $\alpha \mid\sim \beta$ per "α inferisce tentativamente β", e si scrive $\alpha \equiv \beta$ per $\vdash \alpha \leftrightarrow \beta$.

(Reflex) $\alpha \mid\sim \alpha$ (Riflessività: ogni formula è conseguenza non-monotona di se stessa)

(LLE) Se $\alpha \equiv \beta$ e $\alpha \mid\sim \gamma$, allora $\beta \mid\sim \gamma$ (Left Logical Equivalence: la sinistra è chiusa per equivalenza logica)

(RW) Se $\alpha \mid\sim \beta$ e $\vdash \beta \rightarrow \gamma$, allora $\alpha \mid\sim \gamma$ (Right Weakening: le conseguenze logiche di conseguenze non-monotone sono conseguenze non-monotone)

(And) Se $\alpha \mid\sim \beta$ e $\alpha \mid\sim \gamma$, allora $\alpha \mid\sim (\beta \wedge \gamma)$ (Congiunzione a destra: la chiusura per congiunzione delle conseguenze)

(Or) Se $\alpha \mid\sim \gamma$ e $\beta \mid\sim \gamma$, allora $(\alpha \vee \beta) \mid\sim \gamma$ (Disgiunzione a sinistra: se γ segue da α e da β, segue dalla disgiunzione)

(CM) Se $\alpha \mid\sim \beta$ e $\alpha \mid\sim \gamma$, allora $(\alpha \wedge \beta) \mid\sim \gamma$ (Cautious Monotonicity: si può aggiungere una conseguenza alle premesse senza perdere altre conseguenze)

(Cut) Se $\alpha \mid\sim \beta$ e $(\alpha \wedge \beta) \mid\sim \gamma$, allora $\alpha \mid\sim \gamma$ (Cut: eliminazione del lemma nelle inferenze a catena)

I postulati (Reflex), (LLE), (RW), (And) e (Or) definiscono le *logiche cumulative* — il livello minimo di razionalità. (CM) e (Cut) aggiungono il requisito di *cumulatività*, che è la condizione caratteristica delle logiche KLM nel senso stretto. La cumulatività esprime una forma di coerenza locale: ciò che si può inferire da α non cambia se si aggiunge alle premesse una delle conseguenze già inferite.

La semantica che caratterizza questi postulati è quella dei *modelli preferenziali*. Un modello preferenziale è una coppia $(W, <)$ dove W è un insieme di mondi e $<$ è una relazione binaria su W — non necessariamente totale né ben fondata — interpretata come "il mondo w_1 è strettamente preferibile (più normale, più tipico) al mondo w_2". La relazione di conseguenza indotta è:

$\alpha \mid\sim_{<} \beta$ sse ogni mondo $<$-minimale in $[\alpha]$ soddisfa β

ovvero: tipicamente, se α, allora β è vero se in tutti i mondi α-compatibili più normali (più tipici, più plausibili) è vero β. La struttura è formalmente identica alla selezione di Grove per la revisione, con la differenza che qui $<$ non è necessariamente totale.

Il *teorema di rappresentazione KLM* afferma:

Teorema (Kraus–Lehmann–Magidor, 1990). Una relazione $\mid\sim$ soddisfa i postulati (Reflex), (LLE), (RW), (And), (Or), (CM), (Cut) se e soltanto se esiste un modello preferenziale $(W, <)$ con $<$ *smooth* — nel senso che ogni insieme non vuoto ha elementi minimali o ha elementi sotto cui nessun altro elemento è minimale — tale che $\mid\sim = \mid\sim_{<}$.

La condizione di smoothness sostituisce la ben fondatezza richiedendo qualcosa di più debole: non che ogni insieme abbia minimi, ma che ogni elemento non-minimale

abbia sotto di sé elementi più piccoli. Nei modelli finiti le due condizioni coincidono; nei modelli infiniti la smoothness è strettamente più generale.

La dimostrazione della completezza — che ogni relazione $|\sim$ che soddisfi i postulati sia realizzabile da un modello preferenziale — è la parte tecnicamente più ricca. Kraus, Lehmann e Magidor costruiscono il modello canonico come segue. Si definisce su W = {mondi classici} la relazione $<$ ponendo:

$w_1 < w_2$ sse esiste α tale che: $w_1 \models \alpha$, $w_2 \models \alpha$, e $w_2 \notin |\sim_\min([\alpha])$

dove $|\sim min([a])$ *denota i mondi minimali in [a] rispetto all'ordinamento che si sta costruendo. Questa definizione è circolare ma viene risolta mediante un argomento di punto fisso: si definisce* $<$ *come il più piccolo preordine tale che la relazione* $|\sim_<$ che induce soddisfi i postulati, e si verifica che questo punto fisso esiste per il teorema di Knaster–Tarski applicato al reticolo delle relazioni binarie su W.

Movimento XIV — Razionalità e completezza: la condizione di rankability

---

I postulati KLM caratterizzano la razionalità *cumulativa*, ma Lehmann e Magidor riconobbero subito che esiste un livello più forte di razionalità che i modelli preferenziali generali non catturano. La questione emerge considerando un semplice esempio.

Si supponga che l'agente sappia: "i pinguini tipicamente non volano" (Pinguino $|\sim \neg$Vola), "gli uccelli tipicamente volano" (Uccello $|\sim$ Vola), e "i pinguini sono uccelli" ($\vdash$ Pinguino $\rightarrow$ Uccello). Da queste premesse, cosa dovrebbe concludere l'agente di fronte a un pinguino? Un ragionatore pienamente razionale dovrebbe concludere che il pinguino non vola, perché l'informazione "è un pinguino" è più specifica dell'informazione "è un uccello" e dovrebbe prevalere. Ma i postulati KLM non garantiscono questo esito: esistono modelli preferenziali che soddisfano tutte e tre le premesse e che tuttavia non producono Pinguino $|\sim \neg$Vola.

Questo è il famoso problema della *specificità* o della *tipicità differenziata*. Esso mostra che la razionalità cumulativa non è sufficiente a catturare un principio epistemico che sembra ovvio: informazioni più specifiche devono prevalere su informazioni più generali nei conflitti.

La soluzione proposta da Lehmann e Magidor nel 1992 — in un articolo che considerano una continuazione naturale del lavoro KLM — è quella di restringere la classe dei modelli ammissibili ai *modelli rankati*, ovvero ai modelli preferenziali in cui $<$ è un preordine totale: ogni due mondi sono comparabili. In un modello rankato, l'ordinamento di plausibilità divide i mondi in *livelli* o *strati* $\varkappa = 0, 1, 2, ...$, dove $\varkappa = 0$ è il livello dei mondi più normali.

La relazione di conseguenza indotta da un modello rankato è:

$\alpha \mid\sim_\kappa \beta$ sse $\min_\kappa([\alpha]) \subseteq [\beta]$

dove $\min_\kappa([\alpha]) = \{w \in [\alpha] \mid \kappa(w)$ è minimo tra i mondi in $[\alpha]\}$, esattamente la struttura delle sfere di Grove.

Il postulato che distingue i modelli rankati da quelli preferenziali generali è la *razionalità* (Rat), talvolta anche denominata condizione di *rankability*:

(Rat) Se $\alpha \mid\sim \gamma$ e $\neg(\alpha \mid\sim \neg\beta)$, allora $(\alpha \wedge \beta) \mid\sim \gamma$

In parole: se γ segue tipicamente da α, e se β non è *tipicamente escluso* da α (ovvero α non inferisce $\neg\beta$), allora γ segue tipicamente anche da $\alpha \wedge \beta$. Questo cattura l'idea che aggiungere alle premesse qualcosa che non sia tipicamente escluso non deve far perdere conseguenze già inferite — una forma di monotonia condizionale più forte della cautious monotonicity.

Il *teorema di rappresentazione per la razionalità* afferma:

Teorema (Lehmann–Magidor, 1992). Una relazione $\mid\sim$ soddisfa i postulati KLM e (Rat) se e soltanto se esiste un modello preferenziale rankato (W, κ) con $\kappa : W \to \mathrm{Ord}$ (funzione di rango a valori in un ordinale) tale che $\mid\sim = \mid\sim_\kappa$.

La dimostrazione della necessità — che ogni $\mid\sim$ realizzata da un modello rankato soddisfi (Rat) — è immediata dalla definizione: se $w \vDash \alpha \wedge \beta$ e $\kappa(w)$ è minimale in $[\alpha \wedge \beta]$, allora $w \vDash \alpha$ e $\kappa(w) \geq \min_\kappa([\alpha])$. Se $\neg(\alpha \mid\sim \neg\beta)$, esiste $w' \in \min_\kappa([\alpha])$ con $w' \vDash \beta$, dunque $w' \in [\alpha \wedge \beta]$ con $\kappa(w') = \min_\kappa([\alpha]) \leq \kappa(w)$. Poiché $\alpha \mid\sim \gamma$, $w' \vDash \gamma$. Ma w' era arbitrario in $\min_\kappa([\alpha \wedge \beta])$ (per questa catena di disuguaglianze), dunque $\min_\kappa([\alpha \wedge \beta]) \subseteq [\gamma]$, cioè $(\alpha \wedge \beta) \mid\sim_\kappa \gamma$. ✓

La dimostrazione della sufficienza — costruire il modello rankato a partire dai postulati — è più elaborata e procede mediante la nozione di *Z-ranking* introdotta da Pearl. Si definisce una successione di insiemi $Z^0 \supseteq Z^1 \supseteq Z^2 \supseteq \ldots$ di enunciati *tollerabili* rispetto a $\mid\sim$:

$Z^0 = \{\alpha \in L \mid$ esiste β con $\alpha \mid\sim \beta$ e $\neg(\top \mid\sim \neg\alpha)\}$ (enunciati consistenti con la normalità globale)

Z^{i+1} = {α ∈ Z^i | esiste β con α |~ β nella relazione ristretta a Z^i}

Il rango $\kappa(w)$ di un mondo w è il minimo i tale che $w \notin Z^i$ — ovvero il livello di eccezionalità del mondo w rispetto alla gerarchia della normalità. Si verifica che questa costruzione produce un modello rankato che realizza $\mid\sim$, usando (Rat) in modo essenziale per garantire che la gerarchia Z^i sia ben definita e convergente. ✓

Movimento XV — La dualità tra revisione e conseguenza non-monotona

---

A questo punto la connessione strutturale tra la teoria AGM e il framework KLM può essere enunciata con precisione. Entrambe le teorie, nelle loro versioni più raffinate, sono caratterizzate dalla stessa struttura matematica: un preordine totale e ben fondato (o smooth) su W. Le due teorie differiscono soltanto nella *direzione* in cui usano questa struttura.

La revisione AGM usa l'ordinamento $\leqslant_K$ per rispondere alla domanda: dato un corpus K e una nuova informazione α, quale corpus $K * \alpha$ produrre? La risposta è: il corpus corrispondente ai mondi α-minimali secondo $\leqslant_K$.

La conseguenza non-monotona KLM-razionale usa l'ordinamento $\varkappa$ per rispondere alla domanda: data una formula α come premessa, quali formule β ne seguono tipicamente? La risposta è: tutte le β vere nei mondi α-minimali secondo $\varkappa$.

La dualità diventa esplicita attraverso il seguente *teorema di traduzione*, che si trova nella forma più limpida in Makinson e Gärdenfors (1991):

Teorema (Makinson–Gärdenfors, 1991). Esiste una corrispondenza biunivoca tra:

(a) Operazioni di revisione $*$ su corpora che soddisfano (K*1)–(K*8)

(b) Relazioni di conseguenza non-monotona $|\sim$ su L che soddisfano i postulati KLM e (Rat)

La corrispondenza è data dalle mappe:

$\Phi : * \mapsto |\sim*$ *definita da a* $|\sim* \beta$ sse $\beta \in K_\top * \alpha$

$\Psi : |\sim \mapsto *\{|\sim\}$ *definita da K* $*\{|\sim\} \alpha = \{\beta \mid \alpha |\sim^K \beta\}$

dove $K_\top$ è il corpus minimale (le sole tautologie) e $|\sim^K$ è la relazione $|\sim$ ristretta ai mondi compatibili con K. Le due mappe soddisfano $\Psi \circ \Phi = \mathrm{id}$ e $\Phi \circ \Psi = \mathrm{id}$, stabilendo l'isomorfismo.

L'interpretazione filosofica di questo risultato è densa di conseguenze. Essa mostra che la revisione delle credenze e l'inferenza non-monotona non sono due fenomeni distinti che richiedono trattamenti separati, bensì un unico fenomeno descritto da due vocabolari complementari: il vocabolario *doxastico* (corpora, revisione, entrenchment) e il vocabolario *inferenziale* (premesse, conseguenze, tipicità). Un agente che possiede una struttura di plausibilità può usarla sia per aggiornare le proprie credenze sia per inferire tipicità — e la struttura è la stessa.

Questa unificazione ha anche un risvolto critico: le limitazioni dell'uno si rispecchiano nelle limitazioni dell'altro. In particolare, il teorema di impossibilità di Gärdenfors —

che mostra l'incompatibilità tra revisione AGM e semantica condizionale dei condizionali — ha un corrispettivo nel framework KLM: la relazione $|\sim$ non può essere chiusa per *fortalecimento della premessa* (strengthening of the antecedent), ovvero non vale in generale $\alpha \mid\sim \gamma \Longrightarrow (\alpha \wedge \beta) \mid\sim \gamma$. Questo è il prezzo della non-monotonia: si guadagna in potere espressivo rispetto all'inferenza tipica, si perde la monotonia sintattica.

Ma c'è un guadagno più sottile che l'isomorfismo rivela. Nella revisione AGM, il corpus K svolge il ruolo di *contesto epistemico* di sfondo: è ciò che l'agente già sa, e rispetto a cui la nuova informazione viene valutata. Nella conseguenza non-monotona, questo contesto è *implicito* nella struttura dell'ordinamento $\varkappa$. L'isomorfismo rende esplicito ciò che era implicito: l'ordinamento di plausibilità di un agente *è* il suo contesto epistemico di sfondo, e cambiarlo è esattamente ciò che la revisione fa.

Movimento XVI — La logica delle probabilità basse: Adams e le ε-semantiche

---

La connessione tra ragionamento non-monotono e probabilità non è soltanto quella, già esplorata, del condizionamento bayesiano come limite dell'aggiornamento ordinale. Esiste una tradizione parallela — inaugurata da Ernest Adams nei suoi lavori degli anni Sessanta e Settanta — che tratta i condizionali indicativi come enunciati la cui appropriatezza assertiva è governata non dalla verità ma dalla *probabilità condizionale alta*. Questa tradizione, inizialmente sviluppata in modo indipendente dalla teoria AGM, si rivela connessa alla struttura KLM in modo quasi sorprendente.

La proposta fondamentale di Adams è la seguente. Un condizionale indicativo "se α allora β" — da distinguere dal controfattuale sogguntivo di Lewis — non ha condizioni di verità nel senso classico, ma ha condizioni di *assertibilità*: è appropriato asserirlo in un contesto in cui la propria probabilità soggettiva P soddisfa:

$$P(\beta \mid \alpha) \geq 1 - \varepsilon$$

per qualche soglia ε piccola. In altri termini, un condizionale indicativo è asseribile se la probabilità della conseguente dato l'antecedente è alta. Adams non identifica questa alta probabilità condizionale con la verità del condizionale; si tratta di una tesi *pragmatica*, non semantica.

Questa scelta ha una conseguenza formale immediata: la logica dei condizionali di Adams è una logica della *probabilità quasi-nulla dell'eccezione*. Un argomento condizionale è valido — nel senso di Adams — se, ogniqualvolta le premesse sono assertibili (con probabilità degli errori ε_1, ..., ε piccole), la conclusione è assertibile con probabilità dell'errore al massimo $\varepsilon_1 + \ldots + \varepsilon$. Questa nozione di validità è chiamata *p-validità* o *ε-validità*.

Il teorema principale di Adams — dimostrato nella versione più precisa nel suo libro del 1975, *The Logic of Conditionals* — afferma che un argomento condizionale è p-valido

se e soltanto se la sua conclusione appartiene all'insieme delle conseguenze logiche dell'antecedente nel senso delle probabilità che tendono a zero. Formalmente, l'argomento $\{\alpha_1 \to \beta_1, \ldots, \alpha \to \beta\} \vdash \gamma \to \delta$ è p-valido sse: per ogni famiglia di distribuzioni di probabilità P_ε (parametrizzate da $\varepsilon \to 0$) che soddisfino $P_\varepsilon(\beta_i \mid \alpha_i) \to 1$ per ogni i, si ha $P_\varepsilon(\delta \mid \gamma) \to 1$.

Il collegamento con la struttura KLM è ora diretto. Si consideri una famiglia di distribuzioni P_ε costruita a partire da un ranking κ mediante la formula esponenziale già introdotta nel Movimento IX:

$P_\varepsilon(w) = \varepsilon^{\kappa(w)} \cdot Z_\varepsilon^{-1}$

Allora $P_\varepsilon(\beta \mid \alpha) \to 1$ per $\varepsilon \to 0$ sse $\min_\kappa([\alpha]) \subseteq [\beta]$ sse $\alpha \mid\sim_\kappa \beta$. Dunque la p-validità di Adams coincide esattamente con la preservazione della conseguenza non-monotona $\mid\sim_\kappa$ attraverso le deduzioni: un argomento è p-valido sse preserva la relazione $\mid\sim$ in ogni modello rankato.

Teorema (Adams–KLM, versione unificata). Un argomento condizionale è p-valido nel senso di Adams se e soltanto se è valido in ogni modello preferenziale rankato nel senso di Lehmann–Magidor.

Questo risultato — stabilito in modo esplicito da Lehmann e Magidor e riconosciuto da Adams stesso — chiude un cerchio notevole: la logica probabilistica di Adams, sviluppata a partire da intuizioni pragmatiche sull'assertibilità, coincide esattamente con la logica dell'inferenza tipica dei modelli rankati, sviluppata a partire da intuizioni cognitive sulla plausibilità. Due tradizioni intellettualmente distinte convergono sulla stessa struttura matematica.

Movimento XVII — Probabilità qualitativa e il problema della misurabilità

L'identificazione tra p-validità e validità nei modelli rankati suggerisce che la struttura dell'inferenza razionale sotto incertezza sia, in qualche senso fondamentale, *ordinale* piuttosto che cardinale. Questo solleva una domanda che ha radici nella teoria della misurazione: in quali condizioni un ordinamento qualitativo di plausibilità può essere rappresentato da una funzione di probabilità numerica?

La risposta classica è quella della *teoria della probabilità qualitativa*, inaugurata da de Finetti e Koopman e sistematizzata da Scott e Suppes negli anni Cinquanta e Sessanta. Una *relazione di probabilità qualitativa* è una relazione binaria $\preccurlyeq_P$ su $\mathcal{A}$ (una σ-algebra di eventi) che soddisfa:

(QP1) $\preccurlyeq_P$ è un preordine totale su $\mathcal{A}$

(QP2) $\varnothing \preccurlyeq_P \alpha$ per ogni α (l'evento impossibile è il meno probabile)

(QP3) $\alpha \leqslant_P \Omega$ per ogni α (l'evento certo è il più probabile)

(QP4) Se $\alpha \cap \gamma = \varnothing$ e $\beta \cap \gamma = \varnothing$, allora: $\alpha \leqslant_P \beta$ sse $(\alpha \cup \gamma) \leqslant_P (\beta \cup \gamma)$ (Additività qualitativa: aggiungere lo stesso evento disgiunti preserva l'ordine)

Una relazione $\leqslant_P$ che soddisfi (QP1)–(QP4) è detta *relazione di probabilità qualitativa*. La domanda è se esista sempre una funzione di probabilità numerica $P : \mathcal{A} \to [0,1]$ che la rappresenti nel senso che:

$\alpha \leqslant_P \beta$ sse $P(\alpha) \leq P(\beta)$

La risposta — e qui il problema si fa filosoficamente sottile — è *no in generale*. Fine e Gill (1976) e Kraft, Pratt e Seidenberg (1959) hanno costruito esempi di relazioni di probabilità qualitativa su spazi finiti che non ammettono alcuna rappresentazione numerica. Questi esempi mostrano che le condizioni (QP1)–(QP4) sono necessarie ma non sufficienti per la rappresentabilità.

La condizione aggiuntiva necessaria e sufficiente (per spazi finiti) è la cosiddetta *condizione di Scott* o *condizione di coerenza estesa*: per ogni familia finita di eventi $\alpha_1, ..., \alpha$, $\beta_1, ..., \beta \in \mathcal{A}$, se ogni mondo w appartiene ad almeno tanti α_i quanti β_i (ovvero $|\{i \mid w \in \alpha_i\}| \geq |\{i \mid w \in \beta_i\}|$ per ogni w), allora:

non è il caso che $\beta <_P \alpha$ per ogni j

La condizione di Scott, nella sua forma completa, è un schema infinito di condizioni — una per ogni n — e la sua verificazione diretta è computazionalmente intrattabile per n grande. Questo è uno dei motivi per cui la teoria della probabilità qualitativa non ha avuto lo stesso sviluppo applicativo della probabilità numerica: le condizioni di rappresentabilità sono formalmente precise ma pragmaticamente inaccessibili.

Dal punto di vista epistemologico, questa difficoltà ha un significato preciso: un agente che possiede soltanto giudizi comparativi di plausibilità ("questo evento mi sembra più probabile di quello") potrebbe trovarsi in una situazione in cui i propri giudizi, pur localmente coerenti, non sono globalmente rappresentabili da alcuna distribuzione probabilistica. Questo non significa che l'agente sia irrazionale — i postulati (QP1)–(QP4) sono pienamente soddisfatti — ma che la sua struttura epistemica è più ricca (nel senso: più fine-grained) di qualunque distribuzione numerica.

Il legame con gli ordinamenti di entrenchment è diretto. Un ordinamento di entrenchment $\leq_e$ su L può essere interpretato come una relazione di probabilità qualitativa sugli eventi $\{[\alpha] \mid \alpha \in L\}$ se si pone:

$[\alpha] \leqslant_P [\beta]$ sse $\alpha \leq_e \beta$

Ma questa identificazione presuppone che l'entrenchment rispetti (QP4) — la proprietà additiva qualitativa — e la condizione corrispondente per l'entrenchment è precisamente (EE3): la congiuntività. Vi è dunque una corrispondenza strutturale tra

le due teorie, anche se non un'identità: l'entrenchment è definito sull'algebra degli enunciati, la probabilità qualitativa sull'algebra degli eventi.

Movimento XVIII — Il problema della chiusura e il paradosso preface

---

L'assunzione di chiusura logica del corpus K — che K = Cn(K) — è stata il fondamento tecnico dell'intera costruzione. Ma essa è anche il punto dove la teoria si espone alla critica più classica, quella incarnata nel cosiddetto *paradosso del prefazio* di David Makinson (1965) — ironicamente lo stesso Makinson che è coautore della teoria AGM.

L'autore di un libro scrive nel prefazio: "Sono certo che questo libro contenga degli errori". Questa affermazione sembra epistemicamente ragionevole — nessun autore, per quanto scrupoloso, può garantire l'assenza assoluta di errori in un testo di qualche lunghezza. Eppure, se l'autore è razionale nel senso classico, egli crede ogni singola proposizione p_1, p_2, ..., p affermata nel libro — altrimenti non l'avrebbe scritta. Ma allora crede $p_1 \wedge p_2 \wedge \ldots \wedge p$ (per chiusura logica sotto congiunzione), e crede al tempo stesso $\neg(p_1 \wedge p_2 \wedge \ldots \wedge p)$ (perché crede che almeno una delle p_i sia falsa). Il suo corpus è dunque inconsistente.

Il paradosso è particolarmente acuto per la teoria AGM perché attacca proprio (K*5) — la consistenza del corpus — non come effetto di una revisione mal condotta ma come condizione iniziale di qualsiasi agente razionale reale. Un agente con un corpus sufficientemente ricco e una ragionevole consapevolezza della propria fallibilità avrà necessariamente un corpus inconsistente nel senso AGM.

Le risposte al paradosso del prefazio si dividono in due famiglie. La prima famiglia — quella *paraconsistentista* — accetta che il corpus possa essere inconsistente e modifica la logica di sfondo affinché l'inconsistenza non collassi in trivialità. In una logica paraconsistente, da $\alpha \wedge \neg\alpha$ non si deduce β per ogni β (si abbandona il principio ex contradictione quodlibet), e dunque un corpus inconsistente non è automaticamente il corpus triviale $K_{\perp}$. I postulati AGM devono essere ripensati in questo quadro, e il lavoro di Priest e Berto ha mostrato che tale ripensamento è tecnicamente realizzabile, anche se produce strutture notevolmente più complesse.

La seconda famiglia — quella *probabilistica* — sostiene che il paradosso del prefazio mostra semplicemente che la credenza non è binaria: l'autore non crede (nel senso di probabilità 1) ogni singola proposizione del libro, ma la ritiene molto probabilmente vera. La credenza razionale reale è graduata, e la rappresentazione binaria è un'idealizzazione che funziona bene in molti contesti ma che si spezza precisamente dove la fallibilità epistemica diventa rilevante. Il paradosso del prefazio sarebbe dunque non una contraddizione interna alla teoria AGM ma un argomento per il

gradualismo doxastico — la posizione che la credenza propriamente detta sia sempre parziale e mai categorica.

La risposta probabilistica ha però un prezzo: essa dissolve il problema piuttosto che risolverlo. La teoria AGM non pretende di descrivere le credenze parziali — per quello c'è la teoria della probabilità soggettiva — ma le credenze categoriche. La domanda non è se gli agenti reali abbiano credenze graduate (certamente sì) ma se abbia senso parlare di credenze categoriche come oggetto teorico autonomo. Molti epistemologi — tra cui John Pollock, Henry Kyburg e, più recentemente, Hannes Leitgeb — hanno difeso l'autonomia delle credenze categoriche pur ammettendo la loro connessione con quelle graduali.

Il contributo più rilevante su questo tema è la *Tesi della soglia* di Kyburg e la sua riformulazione da parte di Leitgeb come *Tesi della stabilità*. Kyburg proponeva: un agente crede categoricamente α se e soltanto se $P(\alpha) \geq \tau$ per qualche soglia $\tau < 1$. Ma questa proposta produce immediatamente il *paradosso della lotteria*: se $\tau < 1$, l'agente crede di ogni singolo biglietto della lotteria che non vincerà (perché $P(\neg \text{vincerà}_i) = 1 - 1/n > \tau$ per n grande), ma non crede che nessun biglietto vinca — contraddicendo nuovamente la chiusura per congiunzione.

Leitgeb ha proposto nel 2014 una versione più sofisticata: la *Tesi della stabilità* afferma che un insieme B di proposizioni costituisce un corpus categoricamente razionale se e soltanto se B è una proposizione P-stabile, ovvero: $P(\alpha \mid \beta) \geq \tau$ per ogni $\alpha \in B$ e ogni β con $P(\beta) > 0$ che sia compatibile con B. Questa condizione di stabilità condizionale richiede che le credenze in B rimangano alte anche dopo qualsiasi condizionamento compatibile con B — una forma di robustezza epistemica.

Leitgeb dimostra che:

Teorema (Leitgeb, 2014). Per ogni distribuzione di probabilità P, esiste un insieme P-stabile non triviale B se e soltanto se P non è uniforme su un insieme finito di cardinalità ≥ 2. Inoltre, gli insiemi P-stabili con $\tau = 1/2$ sono esattamente le proposizioni che hanno probabilità $> 1/2$ e che sono incluse nella proposizione di probabilità massimale.

Questo risultato ha una conseguenza notevole: per τ abbastanza alto (vicino a 1), il corpus categoricamente razionale di un agente è *unico* dato il suo stato probabilistico, e coincide con l'insieme delle conseguenze della proposizione di probabilità massimale. La struttura è di nuovo quella dei modelli rankati: il corpus categorico è determinato dal livello di plausibilità 0 dell'ordinamento indotto da P.

Il paradosso del prefazio viene così *assorbito* piuttosto che risolto: esso mostra che la razionalità categorica piena è un ideale che richiede una struttura probabilistica molto precisa, e che nella pratica epistemica la tensione tra credenze categoriche e probabilità graduate è ineliminabile. La teoria AGM descrive la struttura di questo ideale con rigore ammirevole; riconoscerne i limiti non è un fallimento ma una specificazione del suo dominio.

Movimento XIX — Chiusura: l'unità formale della mente che rivede se stessa

---

Percorrendo in retrospettiva l'arco di questa monografia, emerge una figura ricorrente: quella di una *struttura ordinale* che si manifesta in vocabolari diversi — le sfere concentriche di Grove, i ranghi di Spohn, i livelli di eccezionalità di Lehmann–Magidor, le soglie di assertibilità di Adams, gli strati di stabilità di Leitgeb — ma che è sempre, nel profondo, la stessa entità matematica: un preordine totale e ben fondato su uno spazio di mondi possibili, ovvero una *gerarchia di plausibilità*.

Questa convergenza non è casuale. Essa riflette un fatto profondo sull'epistemologia normativa: la razionalità di un agente che conosce e che revisiona le proprie conoscenze non può essere catturata soltanto da criteri di *verità* (quali credenze siano vere) né soltanto da criteri di *coerenza* interna (se le credenze siano mutuamente consistenti), ma richiede una struttura di *ordinamento* — una risposta alla domanda "se qualcosa deve essere sacrificato, cosa sacrificare per primo?". Questa struttura è ciò che la tradizione formale chiama plausibilità, entrenchment, ranking, tipicità: nomi diversi per la stessa relazione che un agente intrattiene con i propri stati dossastici.

La filosofia analitica ha spesso cercato di ridurre la nozione di razionalità a criteri algoritmici o computazionali. La teoria AGM e le sue estensioni mostrano che questa riduzione è possibile, ma a un prezzo: ciò che viene catturato formalmente è la struttura *relazionale* della razionalità — come le credenze si connettono tra loro, come cambiano in risposta all'evidenza, come alcune precedono altre nell'ordine del sacrificio — e non il suo *contenuto* — quali credenze siano giustificate dall'evidenza empirica, quale sia l'origine dei giudizi di plausibilità, come l'ordinamento si formi nel tempo.

Questa distinzione tra struttura e contenuto, tra forma e materia della razionalità epistemica, è in ultima analisi la lezione più profonda che la teoria della revisione doxastica consegna alla filosofia della mente. Un agente razionale non è soltanto un insieme di credenze coerenti: è una *struttura ordinata di propensioni al cambiamento*, una gerarchia latente che non si manifesta mai interamente nello stato corrente delle credenze ma che determina, in ogni circostanza futura, come quelle credenze muteranno. La mente che conosce è inseparabile dalla mente che *potrebbe* rivedere se stessa — e la geometria di questi potenziali rivisionamenti è, al tempo stesso, la geometria della razionalità e, in qualche misura, la geometria di ciò che si può chiamare carattere epistemico di un agente.

Appendice A — Strutture algebriche di sfondo

La teoria AGM presuppone una logica di sfondo L che si assume essere la logica proposizionale classica, anche se molti risultati si estendono alla logica del primo ordine con opportune modifiche. In questa appendice si fissano le convenzioni e si enunciano i lemmi tecnici usati senza dimostrazione nel corpo principale.

A.1 Linguaggio e conseguenza. Il linguaggio L è chiuso sotto i connettivi booleani $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$ e contiene una costante $\top$ (tautologia) e $\bot$ (contraddizione). La relazione di conseguenza logica $\vdash$ è quella classica, e Cn : 2^L → 2^L è l'operatore di chiusura per conseguenza: Cn(X) = {α ∈ L | X ⊢ α}. Si verificano le seguenti proprietà standard:

Cn è *estensiva*: X ⊆ Cn(X) Cn è *monotona*: X ⊆ Y ⟹ Cn(X) ⊆ Cn(Y) Cn è *idempotente*: Cn(Cn(X)) = Cn(X) Cn è *compatta*: α ∈ Cn(X) ⟹ α ∈ Cn(Y) per qualche Y ⊆_fin X Cn è *sopraclassica*: se X ⊢_cl α, allora α ∈ Cn(X)

Una *teoria* è un insieme T con Cn(T) = T. L'insieme di tutte le teorie, ordinato per inclusione, forma un reticolo completo in cui l'infimo è l'intersezione e il supremo è Cn(∪·). Il corpus triviale K_⊥ = L è il massimo di questo reticolo.

A.2 Modelli e mondi possibili. Un *mondo possibile* w è una funzione w : Var → {0,1} che assegna valori di verità a tutte le variabili proposizionali di L. L'insieme di tutti i mondi è W = {0,1}^Var. Ogni formula α individua l'insieme dei suoi modelli [α] = {w ∈ W | w ⊨ α}. La mappa α ↦ [α] è un omomorfismo dall'algebra di Lindenbaum di L all'algebra booleana 2^W, ed è iniettiva modulo equivalenza logica.

A.3 Corrispondenza teorie-insiemi di mondi. Esiste una corrispondenza duale tra teorie e insiemi di mondi chiusi:

Teoria T ↦ [T] = ∩_{α ∈ T} [α] (l'insieme dei mondi compatibili con T) Insieme M ↦ Th(M) = {α | M ⊆ [α]} (la teoria degli enunciati veri in tutti i mondi di M)

Le mappe soddisfano [Th(M)] ⊇ M e Th([T]) = T per ogni teoria T. La corrispondenza è un'isomorfismo tra il reticolo delle teorie (ordinato per ⊆) e il reticolo degli insiemi di mondi chiusi per Th (ordinato per ⊇ — l'ordine è dualizzato).

Appendice B — Preordini, buoni ordini e sistemi di sfere: definizioni complete

B.1 Preordini e ordini. Una relazione binaria $\preccurlyeq$ su un insieme X è un *preordine* se è riflessiva (x $\preccurlyeq$ x per ogni x) e transitiva (x $\preccurlyeq$ y e y $\preccurlyeq$ z implicano x $\preccurlyeq$ z). È un *ordine parziale* se è anche antisimmetrica (x $\preccurlyeq$ y e y $\preccurlyeq$ x implicano x = y). È *totale* (o lineare) se per ogni x, y vale x $\preccurlyeq$ y oppure y $\preccurlyeq$ x. È *ben fondata* se ogni sottoinsieme non vuoto di X ha un elemento minimale rispetto a $\preccurlyeq$.

Per un preordine $\preccurlyeq$ su X, la relazione di equivalenza associata è $x \sim y$ sse $x \preccurlyeq y$ e $y \preccurlyeq x$. Il quoziente $X/\sim$ è un ordine parziale (o totale se $\preccurlyeq$ era totale). Gli elementi minimali di un sottoinsieme $Y \subseteq X$ rispetto a $\preccurlyeq$ sono:

$$\min_{\preccurlyeq}(Y) = \{y \in Y \mid \nexists\, z \in Y : z \prec y\}$$

dove $z \prec y$ significa $z \preccurlyeq y$ e $\neg(y \preccurlyeq z)$. La ben fondatezza garantisce $\min_{\preccurlyeq}(Y) \neq \varnothing$ per ogni $Y \neq \varnothing$.

B.2 Sistemi di sfere. Riprendiamo la definizione completa. Un *sistema di sfere* su W centrato su un insieme $C \subseteq W$ è una famiglia $\$ \subseteq 2^W$ tale che:

(S1) $\$$ è totalmente ordinata da $\subseteq$ (S2) $C \in \$$ (S3) $C = \bigcap \$$ (C è l'elemento minimo di $\$$) (S4) $W \in \$$

Dato α con $[\alpha] \neq \varnothing$, la *sfera minima per a* è $S_\alpha = \bigcap\{A \in \$ \mid A \cap [\alpha] \neq \varnothing\}$. Poiché $\$$ è totalmente ordinata, l'intersezione di una famiglia di sfere che intersecano $[\alpha]$ è ancora una sfera (o un limite di sfere), e S_α è ben definita. Se nessuna sfera finita interseca $[\alpha]$ si ha $S_\alpha = W$.

La corrispondenza tra sistemi di sfere e preordini totali su W è data da:

Sistema di sfere $\$ \longrightarrow$ Preordine $\preccurlyeq_\$$: $w_1 \preccurlyeq_\$ w_2$ sse $\forall A \in \$: w_1 \in A \Rightarrow w_2 \in A$

Preordine $\preccurlyeq \longrightarrow$ Sistema di sfere $\$_\preccurlyeq = \{\, \{w \mid \varkappa(w) \leq i\} \mid i \in \mathrm{Ord} \,\}$

dove $\varkappa(w) = \min\{i \mid w \text{ è nel livello } i \text{ di } \preccurlyeq\}$ è il rango ordinale di w. Le due costruzioni sono inverse l'una dell'altra a meno di isomorfismo.

Appendice C — Dimostrazione del teorema di Grove: direzione da postulati a sfere

Questo è il cuore tecnico più delicato del Movimento III. Sia $*$ un'operazione che soddisfa (K*1)–(K*8). Si definisce il preordine di plausibilità su W come segue:

Definizione C.1. Per $w_1, w_2 \in W$, si pone:

$$w_1 \preccurlyeq_* w_2 \text{ sse per ogni } \alpha \in L\text{: se } w_2 \in [K * \alpha] \text{ allora } w_1 \in [K * \alpha]$$

Lemma C.2 (Totalità). $\preccurlyeq_*$ è un preordine totale su W.

Dimostrazione. La riflessività è immediata. Per la transitività: se $w_1 \preccurlyeq_* w_2$ *e* $w_2 \preccurlyeq_* w_3$, e $w_3 \in [K * \alpha]$, allora $w_2 \in [K * \alpha]$ (per la seconda condizione), dunque $w_1 \in [K * \alpha]$ (per la prima). La totalità: si supponga per assurdo che né $w_1 \preccurlyeq_* w_2$ *né* $w_2 \preccurlyeq_* w_1$. Allora esistono α, β tali che $w_2 \in [K * \alpha]$ ma $w_1 \notin [K * \alpha]$, e $w_1 \in [K * \beta]$ ma $w_2 \notin [K * \beta]$. Si consideri $\gamma = \mathrm{Th}(\{w_1, w_2\})$ — la formula vera esattamente nei mondi w_1 e w_2 (che

esiste nel linguaggio se W è finito, o come limite nel caso generale). Per (K*2), $K * \gamma$ contiene γ, dunque $[K * \gamma] \subseteq [\gamma] = \{w_1, w_2\}$ per (K*3) e inclusione. Ma $[K * \gamma] \neq \varnothing$ per (K*5). Dunque $w_1 \in [K * \gamma]$ oppure $w_2 \in [K * \gamma]$ (o entrambi). Se $w_1 \in [K * \gamma]$, allora per definizione $w_1 \leqslant_* w_2$ (poiché $w_2 \in [K * \gamma] \Rightarrow w_1 \in [K * \gamma]$), contraddizione. Il caso $w_2 \in [K * \gamma]$ è simmetrico. ∎

Lemma C.3 (I mondi di [K] sono minimi). Per ogni $w \in [K]$ e ogni $v \in W$, si ha $w \leqslant_* v$.

Dimostrazione. Sia $v \in [K * \alpha]$ per qualche α. Per (K*4), se $\neg\alpha \notin K$, allora $K * \alpha = K + \alpha$, dunque $[K * \alpha] \supseteq [K] \cap [\alpha]$. Se $w \in [K]$ e $w \in [\alpha]$, allora $w \in [K * \alpha]$. Se $w \notin [\alpha]$, la situazione richiede (K*3): $[K * \alpha] \subseteq [K + \alpha] = [K] \cap [\alpha]$ non include w, ma questo non è un problema poiché vogliamo mostrare che $w \leqslant_* v$ ogniqualvolta $v \in [K * \alpha]$. Si verifica separatamente per il caso $\neg\alpha \in K$ usando (K*5). ∎

Teorema C.4 (Ricostruzione). Sia S il sistema di sfere indotto da $\leqslant_*$. Allora $[K * \alpha] = S_\alpha \cap [\alpha]$ per ogni α consistente.

Dimostrazione. ($\subseteq$) Sia $w \in [K * \alpha]$. Per (K*2), $w \in [\alpha]$. Si deve mostrare che $w \in S_\alpha$, ovvero che w appartiene alla sfera minima che interseca $[\alpha]$. Poiché $w \in [K * \alpha]$ e $[K * \alpha] \cap [\alpha] \neq \varnothing$ (per (K*2)), la sfera S_α è quella determinata dai mondi di $[K * \alpha]$. Per costruzione del preordine, ogni $v \in [K * \alpha]$ soddisfa $v \leqslant_* w$ per ogni $w \notin [K * \alpha]$, dunque i mondi di $[K * \alpha]$ sono esattamente quelli di $S_\alpha \cap [\alpha]$.

($\supseteq$) Sia $w \in S_\alpha \cap [\alpha]$ ma $w \notin [K * \alpha]$. Allora esiste $v \in [K * \alpha]$ con $v <_* w$ *(v è strettamente più plausibile). Ma* $v \in [K * \alpha]$ *e* $w \notin [K * \alpha]$ *implica per definizione di* $\leqslant_*$ che non vale $w \leqslant_* v$, *cioè* $w <_* v$ o v e w sono incomparabili. Poiché $\leqslant_*$ *è totale,* $v <_* w$, contraddicendo la supposizione che $w \in S_\alpha$ (i mondi di S_α sono quelli di plausibilità minima in $[\alpha]$). ∎

Appendice D — La costruzione Z-ranking di Pearl: dettaglio algoritmico

La costruzione del ranking $\varkappa$ a partire da una relazione $|\sim$ che soddisfa i postulati KLM e (Rat) procede mediante la seguente procedura iterativa, dovuta a Pearl (1990) nella sua trattazione dei *sistemic entailment* e ripresa da Lehmann–Magidor (1992).

Input: Una relazione di conseguenza non-monotona $|\sim$ su L che soddisfa (Reflex), (LLE), (RW), (And), (Or), (CM), (Cut), (Rat).

Passo 0. Si definisce la *base condizionale* $\Delta = \{(\alpha, \beta) \mid \alpha \mid\sim \beta \text{ e } \neg\vdash \neg\alpha\}$, ovvero l'insieme di tutte le coppie di inferenza tipica con antecedente consistente.

Passo 1. Si definisce per ogni formula α la sua *tolleranza*: α è tollerabile (rispetto a Δ) se esiste un mondo $w \vDash \alpha$ che non viola alcun condizionale in Δ in modo *materiale*, ovvero

tale che per ogni $(\gamma, \delta) \in \Delta$ con $w \models \gamma$, si abbia $w \models \delta$. L'insieme dei formula tollerate è:

$$\mathrm{Tol}(\Delta) = \{\alpha \in L \mid \exists\, w \models \alpha : \forall\, (\gamma,\delta) \in \Delta,\, w \models \gamma \Rightarrow w \models \delta\}$$

Passo 2 (iterativo). Si costruisce la successione:

$$\Delta_0 = \Delta \quad \Delta_{i+1} = \{(\alpha, \beta) \in \Delta_i \mid \alpha \in \mathrm{Tol}(\Delta_i)\}$$

ovvero ad ogni passo si rimuovono da Δ_i tutti i condizionali il cui antecedente è *intollerabile* rispetto al sottoinsieme corrente. La successione $\Delta_0 \supseteq \Delta_1 \supseteq \Delta_2 \supseteq \ldots$ è decrescente e, su L finito, si stabilizza in un numero finito di passi. Si chiama Δ_∞ il suo limite.

Assegnazione del rango. Per ogni mondo w, il suo rango $\kappa(w)$ è:

$$\kappa(w) = \min\{i \mid w \in \mathrm{Tol}(\Delta_i)\} \text{ se tale } i \text{ esiste, altrimenti } \kappa(w) = \infty$$

I mondi con $\kappa(w) = 0$ sono quelli *completamente normali* — soddisfano tutti i condizionali. I mondi con $\kappa(w) = i$ sono quelli che cominciano a essere tollerabili solo quando si rimuovono i condizionali dell'i-esimo strato.

Verifica. Si dimostra che κ è una funzione di rango che induce una relazione $|\sim_\kappa$ coincidente con $|\sim$, usando (Rat) per garantire che $\mathrm{Tol}(\Delta_i) \neq \varnothing$ per ogni i tale che $\Delta_i \neq \varnothing$ — questo è il punto dove (Rat) è indispensabile. Senza (Rat), la costruzione potrebbe produrre strati vuoti e quindi un rango non ben definito.

La complessità computazionale della costruzione Z-ranking è polinomiale nel numero di condizionali in Δ per ogni passo dell'iterazione, e il numero di passi è al più $|L|$ (la cardinalità del linguaggio), dunque la complessità totale è al più quadratica in $|\Delta| \cdot |L|$.

Appendice E — Postulati AGM per la contrazione: enunciazione completa

Per completezza, si enunciano i sei postulati di base e i due supplementari per la contrazione. L'operazione $K - \alpha$ deve soddisfare:

(K−1) $K - \alpha = \mathrm{Cn}(K - \alpha)$ (chiusura)

(K−2) $K - \alpha \subseteq K$ (inclusione: la contrazione non aggiunge)

(K−3) Se $\alpha \notin K$, allora $K - \alpha = K$ (vacuità: nulla da rimuovere se α non c'è)

(K−4) Se $\nvdash \alpha$, allora $\alpha \notin K - \alpha$ (successo: la contrazione riesce sui non-tautologie)

(K−5) $K \subseteq (K - \alpha) + \alpha$ (recupero: ciò che si rimuove può essere recuperato espandendo)

(K−6) Se $\vdash \alpha \leftrightarrow \beta$, allora $K - \alpha = K - \beta$ (estensionalità)

(K−7) $(K - \alpha) \cap (K - \beta) \subseteq K - (\alpha \wedge \beta)$ (congiuntività I)

(K−8) Se $\alpha \notin K - (\alpha \wedge \beta)$, allora $K - (\alpha \wedge \beta) \subseteq K - \alpha$ (congiuntività II)

Il postulato di recupero (K−5) è il più controverso: afferma che espandendo il corpus contratto con la stessa formula rimossa si recupera l'intero corpus originario. Levi ha argomentato che questo postulato è eccessivamente forte: esso impone che la contrazione non distrugga informazioni ma le renda soltanto latenti, il che presuppone una memoria della struttura originaria che non è sempre epistemicamente giustificata.

La connessione con l'espansione e la revisione attraverso le identità di Levi e Harper è:

$K * \alpha = (K - \neg\alpha) + \alpha$ (identità di Levi) $K - \alpha = K \cap (K * \neg\alpha)$ (identità di Harper)

Si verifica che se $*$ soddisfa (K*1)–(K*8) e si definisce $-$ tramite l'identità di Harper, allora $-$ soddisfa (K−1)–(K−8), e viceversa.

Appendice F — Lemma di compattezza per i sistemi di sfere infiniti

Nei sistemi di sfere su spazi di mondi infiniti, la definizione di S_α come intersezione di sfere che intersecano $[\alpha]$ può produrre insiemi vuoti se la famiglia è decrescente senza limite inferiore. Il lemma seguente garantisce che questo problema non si presenti sotto ipotesi naturali.

Lemma F.1 (Compattezza delle sfere). Sia S un sistema di sfere su W che sia *compatto* nel senso che ogni famiglia filtrante decrescente $\{A_i\}_{i \in I} \subseteq S$ con $\bigcap_i (A_i \cap [\alpha]) \neq \varnothing$ per ogni sottoinsieme finito ha $\bigcap_i A_i \cap [\alpha] \neq \varnothing$. Allora $S_\alpha \cap [\alpha] \neq \varnothing$ per ogni α consistente.

Dimostrazione. Sia $\mathcal{F} = \{A \in S \mid A \cap [\alpha] \neq \varnothing\}$. $\mathcal{F}$ è un filtro su S rispetto a $\subseteq$. L'ipotesi di compattezza applicata a $\mathcal{F}$ garantisce che $\bigcap \mathcal{F} \cap [\alpha] = S_\alpha \cap [\alpha] \neq \varnothing$. ∎

Questo lemma è la ragione per cui, nella semantica di Grove, si assume solitamente che W sia finito o che il preordine $\preccurlyeq$ sia ben fondato: entrambe le ipotesi garantiscono la compattezza in senso rilevante e quindi la non-vacuità di $S_\alpha \cap [\alpha]$ per ogni α consistente, che è il presupposto del postulato (K*5).

Bibliografia ragionata

Le voci sono organizzate per nuclei tematici anziché alfabeticamente, al fine di rendere esplicite le linee di sviluppo intellettuale tracciate nella monografia. Entro ciascun nucleo l'ordine è cronologico.

I. Fondamenti della teoria AGM

Alchourrón, Carlos E., Peter Gärdenfors e David Makinson — "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", *The Journal of Symbolic Logic*, vol. 50, n. 2, 1985, pp. 510–530. È il testo fondativo. Introduce i postulati, le contrazioni partial meet e l'equivalenza tra contrazione e revisione attraverso le identità di Levi e Harper. La dimostrazione del teorema di rappresentazione per la contrazione partial meet è svolta con rigore completo, e l'appendice sulle funzioni di selezione basate sull'entrenchment anticipa il lavoro di Gärdenfors e Makinson del 1988.

Gärdenfors, Peter — *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge (MA), 1988. Monografia sistematica che raccoglie e organizza i risultati del gruppo AGM, aggiungendo estese riflessioni filosofiche sulla natura della credenza e della revisione. I capitoli sulla connessione tra revisione e condizionali sogguntivi contengono la versione più elaborata dell'argomento che porta al teorema di impossibilità. Rimane la presentazione più accessibile e completa della teoria classica.

Makinson, David — "How to Give It Up: A Survey of Some Formal Aspects of the Logic of Theory Change", *Synthese*, vol. 62, n. 3, 1985, pp. 347–363. Articolo panoramico scritto contestualmente ad AGM (1985) che tratta il problema del recupero e le difficoltà della contrazione con una franchezza notevole. Makinson riconosce qui, ante litteram, alcune delle tensioni che diventeranno oggetto di dibattito negli anni successivi.

Gärdenfors, Peter e David Makinson — "Revisions of Knowledge Systems Using Epistemic Entrenchment", in *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge* (TARK II), a cura di Moshe Vardi, Morgan Kaufmann, San Mateo (CA), 1988, pp. 83–95. Introduce gli ordinamenti di entrenchment epistemico e dimostra il teorema di rappresentazione duale rispetto a quello di Grove. Contiene la costruzione dell'entrenchment a partire dall'operazione di revisione, che è la parte più originale.

II. Semantica delle sfere e modelli preferenziali

Grove, Adam — "Two Modellings for Theory Change", *Journal of Philosophical Logic*, vol. 17, n. 2, 1988, pp. 157–170. Introduce la semantica delle sfere e dimostra la sua equivalenza con i postulati AGM. La costruzione del preordine canonico a partire dall'operazione di revisione è l'argomento tecnico centrale. L'articolo introduce anche il raccordo con le distribuzioni probabilistiche esponenziali, anticipando il filone di Spohn–Halpern.

Lewis, David K. — *Counterfactuals*, Harvard University Press, Cambridge (MA), 1973. Non è un testo di teoria della revisione, ma è la fonte delle strutture a sfere e degli ordinamenti di prossimità tra mondi che Grove eredita direttamente. La semantica lewisiana dei condizionali controfattuali è il precedente formale più diretto per l'intera tradizione della revisione basata su mondi possibili.

Stalnaker, Robert — "A Theory of Conditionals", in *Studies in Logical Theory*, a cura di Nicholas Rescher, Blackwell, Oxford, 1968, pp. 98–112. Insieme a Lewis, è il fondamento della semantica condizionale tramite selezione di mondi. Il modello di Stalnaker differisce da quello di Lewis per l'assunzione che la funzione di selezione sia una funzione (non una relazione), il che produce una semantica deterministica rispetto a quella lewisiana.

III. Il teorema di impossibilità e i condizionali

Gärdenfors, Peter — "Belief Revision and the Ramsey Test for Conditionals", *The Philosophical Review*, vol. 95, n. 1, 1986, pp. 81–93. Contiene la dimostrazione del teorema di impossibilità nella sua forma originale. L'argomento usa i postulati (K*3) e (K*4) insieme al principio di Ramsey per derivare la contraddizione. È uno dei rari casi in cui un risultato tecnico di logica epistemica ha avuto una risonanza significativa nella filosofia analitica mainstream.

Levi, Isaac — *The Fixation of Belief and Its Undoing*, Cambridge University Press, Cambridge, 1991. Critica sistematica della teoria AGM dal punto di vista pragmatista e decisionale. Levi rifiuta l'identità di Levi (che considera un'inversione dell'ordine logico delle operazioni), propone un trattamento autonomo della contrazione come deliberazione, e argomenta che il teorema di impossibilità di Gärdenfors è un argomento contro il principio di Ramsey piuttosto che contro la revisione AGM.

Rott, Hans — *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*, Oxford Logic Guides 42, Oxford University Press, Oxford, 2001. Trattazione enciclopedica che copre sia la teoria AGM sia il ragionamento non-monotono, con attenzione particolare alle connessioni formali tra i due campi. Contiene le dimostrazioni più complete disponibili nella letteratura secondaria per numerosi risultati citati nella monografia.

IV. Revisione iterata

Darwiche, Adnan e Judea Pearl — "On the Logic of Iterated Belief Revision", *Artificial Intelligence*, vol. 89, n. 1–2, 1997, pp. 1–29. Introduce i quattro postulati (C1)–(C4) per la coerenza della revisione iterata e dimostra il teorema di rappresentazione corrispondente. La discussione degli esempi motivanti — in particolare l'esempio del dado e quello degli uccelli-pinguini in versione doxastica — è particolarmente illuminante.

Boutilier, Craig — "Iterated Revision and Minimal Change of Conditional Beliefs", *Journal of Philosophical Logic*, vol. 25, n. 3, 1996, pp. 263–305. Introduce la revisione naturale come alternativa alla revisione lessicografica, con una caratterizzazione assiomatica della politica di aggiornamento che minimizza il cambiamento strutturale dell'ordinamento di plausibilità. Precedente a Darwiche–Pearl ma convergente nei risultati principali.

Nayak, Abhaya C. — "Iterated Belief Change Based on Epistemic Entrenchment", *Erkenntnis*, vol. 41, n. 3, 1994, pp. 353–390. Tratta la revisione iterata dal lato dell'entrenchment anziché degli ordinamenti di plausibilità, ottenendo una caratterizzazione equivalente ma con strumenti distinti. Utile per chi preferisce il vocabolario endogeno.

V. Ragionamento non-monotono e logiche KLM

Kraus, Sarit, Daniel Lehmann e Menachem Magidor — "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics", *Artificial Intelligence*, vol. 44, n. 1–2, 1990, pp. 167–207. È il testo fondativo del framework KLM. I postulati per le logiche cumulative, la semantica preferenziale, e il teorema di rappresentazione con la condizione di smoothness sono tutti introdotti qui. La costruzione del modello canonico per punto fisso è svolta in dettaglio nell'appendice tecnica.

Lehmann, Daniel e Menachem Magidor — "What Does a Conditional Knowledge Base Entail?", *Artificial Intelligence*, vol. 55, n. 1, 1992, pp. 1–60. La continuazione naturale di KLM (1990). Introduce la condizione di razionalità (Rat), i modelli rankati, e il Z-ranking di Pearl come strumento di costruzione del modello canonico. Contiene anche la dimostrazione del teorema di Adams–KLM che unifica le due tradizioni.

Makinson, David e Peter Gärdenfors — "Relations Between the Logic of Theory Change and Nonmonotonic Logic", in *The Logic of Theory Change*, a cura di André Fuhrmann e Michael Morreau, Lecture Notes in Computer Science 465, Springer, Berlin, 1991, pp. 185–205. Stabilisce formalmente l'isomorfismo tra revisione AGM e conseguenza non-monotona razionale. Contiene le mappe Φ e Ψ e la verifica della loro biettività. È il testo che più esplicitamente mostra l'unità formale dei due campi.

Pearl, Judea — "System Z: A Natural Ordering of Defaults with Tractable Applications to Nonmonotonic Reasoning", in *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge* (TARK III), Morgan Kaufmann, 1990, pp. 121–135. Introduce il System Z e la costruzione iterativa degli strati di tolleranza che

diventerà il Z-ranking nella trattazione di Lehmann–Magidor. Contiene anche una discussione dell'esempio del pinguino che ha reso celebre il problema della specificità.

VI. Connessioni probabilistiche e condizionali di Adams

Adams, Ernest W. — *The Logic of Conditionals: An Application of Probability to Deductive Logic*, Reidel, Dordrecht, 1975. L'opera sistematica di Adams sulla logica probabilistica dei condizionali indicativi. Il teorema di p-validità e la sua connessione con le probabilità che tendono ai valori estremi sono qui dimostrati nella forma più completa. La parte filosofica, sulla distinzione tra condizionali indicativi e sogguntivi, rimane influente nel dibattito semantico contemporaneo.

Spohn, Wolfgang — *The Laws of Belief: Ranking Theory and Its Philosophical Applications*, Oxford University Press, Oxford, 2012. La presentazione sistematica e definitiva della teoria dei ranking di Spohn, sviluppata in forma embrionale già nei primi anni Ottanta ma qui portata alla piena maturità. Il Capitolo 5 contiene la regola di aggiornamento per le funzioni di ranking e la sua connessione con la revisione AGM; il Capitolo 11 tratta la relazione con la probabilità soggettiva.

Halpern, Joseph Y. — *Reasoning About Uncertainty*, MIT Press, Cambridge (MA), 2003. Trattazione unificata di tutti i formalismi per ragionare sotto incertezza: probabilità, insiemi di probabilità, funzioni di possibilità, ranking, preordini di plausibilità. Il Capitolo 3 contiene la dimostrazione più rigorosa disponibile della convergenza delle distribuzioni esponenziali verso gli ordinamenti di plausibilità.

van Fraassen, Bas C. — "Probabilities of Conditionals", in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, vol. I, a cura di William Harper e Clifford Hooker, Reidel, Dordrecht, 1976, pp. 261–308. Contiene il risultato noto come "triviality theorem" di Lewis–van Fraassen: se si richiede che $P(\alpha \to \beta) = P(\beta \mid \alpha)$ per ogni α, β e per tutte le distribuzioni di probabilità aggiornabili per condizionamento, allora la logica dei condizionali collassa sulla logica materiale. Complementare al teorema di impossibilità di Gärdenfors.

VII. Probabilità qualitativa, credenza categorica e paradossi epistemici

Kraft, Charles H., John W. Pratt e A. Seidenberg — "Intuitive Probability on Finite Sets", *The Annals of Mathematical Statistics*, vol. 30, n. 2, 1959, pp. 408–419. Contiene la costruzione originale di ordinamenti qualitativi di probabilità su spazi finiti che non ammettono rappresentazione numerica. Il controesempio è di cardinalità 5, il minimo possibile per cui il fenomeno si manifesta.

Scott, Dana — "Measurement Structures and Linear Inequalities", *Journal of Mathematical Psychology*, vol. 1, n. 2, 1964, pp. 233–247. Enuncia la condizione di Scott

come condizione necessaria e sufficiente per la rappresentabilità di una relazione qualitativa di probabilità da parte di una misura numerica. La dimostrazione usa tecniche di programmazione lineare (il duale del teorema di Farkas).

Makinson, David — "The Paradox of the Preface", *Analysis*, vol. 25, n. 6, 1965, pp. 205–207. Il breve articolo originale — meno di tre pagine — in cui Makinson enuncia il paradosso. La sua concisione è inversamente proporzionale alla profondità delle questioni che apre, che riguardano il rapporto tra chiusura logica, fallibilismo epistemico e razionalità delle credenze aggregate.

Leitgeb, Hannes — "The Stability Theory of Belief", *The Philosophical Review*, vol. 123, n. 2, 2014, pp. 131–171. Introduce la Tesi della Stabilità come soluzione al problema della relazione tra credenza categorica e probabilità graduata, e dimostra il teorema di unicità e coincidenza con la struttura di plausibilità 0. È il contributo più originale degli ultimi anni al problema dell'integrazione tra le due tradizioni dossastiche.

Kyburg, Henry E. — *Probability and the Logic of Rational Belief*, Wesleyan University Press, Middletown (CT), 1961. La fonte della Tesi della Soglia originale e del paradosso della lotteria che essa genera. Kyburg difende con coerenza una posizione che accetta le conseguenze controverse della soglia, inclusa la rinuncia alla chiusura per congiunzione delle credenze categoriche.

VIII. Letteratura filosofica di contesto

Ramsey, Frank P. — "General Propositions and Causality" (1929), in *Foundations of Mathematics and Other Logical Essays*, a cura di Richard B. Braithwaite, Routledge, London, 1931, pp. 237–255. Contiene il frammento in cui Ramsey suggerisce il test per i condizionali che porta il suo nome. Il testo è postumo e frammentario, e la ricostruzione del suo pensiero esatto è stata oggetto di discussione.

Brentano, Franz — *Psychologie vom empirischen Standpunkte* (1874), tr. it. *Psicologia dal punto di vista empirico*, Laterza, Bari, 1997. Citato non per i suoi contributi diretti alla teoria della revisione — che sarebbero anacronistici — ma perché la nozione di intenzionalità che Brentano reintroduce nella filosofia moderna è il contesto in cui si inscrive qualsiasi teoria formale della credenza: la credenza è intentional in quanto è sempre credenza *che* qualcosa è il caso, ovvero ha struttura proposizionale. I postulati AGM sono, in un senso preciso, una teoria della razionalità degli stati intenzionali proposizionali.

Quine, Willard Van Orman — "Two Dogmas of Empiricism", *The Philosophical Review*, vol. 60, n. 1, 1951, pp. 20–43. Il testo quineano è rilevante perché la nozione di *entrenchment* epistemico è, nei suoi contorni informali, già presente nella metafora quineana della rete di credenze: alcune credenze sono al centro della rete e vengono modificate soltanto in extremis, altre sono alla periferia e vengono revisionate più facilmente. Gärdenfors stesso riconosce questo debito.

Williamson, Timothy — *Knowledge and Its Limits*, Oxford University Press, Oxford, 2000. La difesa del *knowledge-first approach* all'epistemologia è rilevante per la teoria AGM perché Williamson argomenta che la credenza non è il concetto epistemico primitivo ma lo è la *conoscenza*. Se questa tesi fosse corretta, una teoria della revisione delle credenze che non incorpori la distinzione tra credenza vera giustificata e conoscenza sarebbe fondamentalmente incompleta. Williamson non tratta AGM direttamente, ma il confronto implicito è produttivo.